ConSOLE XXV:

Proceedings of the 25th

Conference of the Student

Organization of Linguistics

in Europe

(4-6 January 2017, University
of Leipzig, Germany)

Edited by

Kate Bellamy

Anastasiia Ionova

George Saad

The syntax and semantics of event measurements in Mandarin

Anqi Zhang

In this paper, I explore the syntax and semantics of the prenominal duration and frequency phrases in Mandarin. I offer an explicit syntactic analysis of the 'verb+ duration/frequency phrase+ direct object' construction in Mandarin Chinese, and argue that the duration/frequency phrase forms a constituent with the direct object in this particular construction. Based on this syntactic account, I propose that the semantics of the direct object has a shifted event description meaning that can be measured directly by the duration phrase and counted by the frequency phrase. This analysis supports Kennedy's (2012) semantic analysis of incremental change.

1.   Introduction

The telicity of an incremental change verb depends on the referential properties of its internal arguments (Krifka 1989, 1998; Tenny 1992; Piñón 2008). However, the compositional semantics of the incremental theme has been mostly analyzed differently for various subtypes, such as consumption verbs, and degree achievements. For a unified account, Kennedy (2012) proposes to extend the scalar analysis of degree achievements to incremental consumption verbs, such that the direct object denotes a measure of change function as shown in (1). However, this type of semantics is not expected of noun phrases, because normally noun phrases are not associated with an event variable.

(1) $[\![\text{dumpling(s)}_{inc}]\!] = \lambda d\lambda x\lambda e.dumplings(x) \wedge NU_{\Delta}(dumplings)(x)(e) = d$

To support Kennedy's analysis, it is crucial to show that the theme NP can be indeed associated with an event variable. In this paper, I argue that the theme direct object in Mandarin can indeed have an event type of meaning in a particular construction of the form 'verb + duration/ frequency phrase + NP'. I show that because the direct object forms a single constituent with the preceding duration phrase or frequency phrase in syntax, semantically the direct object as the

theme must have an event type of meaning that can be measured out by these event measurement phrases.

As shown in (2), in this particular construction, the duration phrase or the frequency phrase appears after the verb and immediately before the direct object. I will refer to this construction as the prenominal duration or frequency phrase construction in the rest of the paper.

(2)   wo tiao-le   san   xiaoshi/ci wu.
        I   dance-PFV three hour/time dance.
        'I danced for three hours/three times.'

This construction raises several interesting theoretical questions (Liao 2014). As noted by Huang et al. (2009), the prenominal duration phrase presents a case of a syntactic-semantic mismatch. For example, as shown in (3), the prenominal duration phrase can be followed by an optional modifier head *de*, which serves as an adjective marker or a genitive marker. In other words, instead of modifying the verb, the duration phrase is unexpectedly modifying the direct object. Because the direct object 'book' is usually not modifiable by a duration phrase 'three hours', given that a book does not have an inherent duration of time, it is puzzling why the duration phrase can form a constituent with the direct object in this construction.

(3)   Ta   du le   san   xiaoshi (de)   shu
        S/he read PFV three hour   (MOD) book
        'S/he read books (or a book) for three hours'

Although in this specific construction, the duration phrase with *de* has been taken to be a single constituent with the direct object, it is unclear whether the prenominal duration phrase and the prenominal frequency phrase without *de* should be analyzed similarly. In fact, the prenominal duration phrase and the prenominal frequency have been predominantly analyzed as not forming a single constituent with the direct object. They have been predominantly analyzed either as special postverbal verbal adjuncts (Huang et al. 2009), or noun arguments of the verb (Cheng et al. 1997). These two analyses seem intuitive at first glance, because they solve the syntactic-semantic mismatch problem by still associating the duration and the frequency phrase with the verb through movement. However, these two analyses lack substantial syntactic evidence for their specific claims. A third lesser known but promising analysis is that the prenominal duration or frequency phrase forms a single constituent with the following direct object, more specifically a 'numeral-classifier' phrase (Sybesma 1999; Liao 2014). According to Liao (2014), the prenominal duration phrase can be shown to form a constituent with the following object with evidence from the tonal sandhi behavior in Taiwanese Chinese. However, evidence from tonal sandhi alone seems to be not strong enough to prove the case, and Liao (2014) unfortunately also does spell out the structures fully.

Extending on Sybesma's (1999) and Liao's (2014) 'numeral+classifier' proposal, I claim that indeed both the prenominal duration and frequency phrases form a constituent with the direct object, but they are structurally different and parallel to different types of classifiers. Crucially, filling a gap in the previous studies, I present direct evidence from constituency tests to show that these constructions form a constituent of their own. Furthermore, I add to Sybesma's discussion of the parallelism between the prenominal duration or frequency phrases and different types of classifiers comparing the two in a new way to show that the prenominal duration and the prenominal frequency phrases are like measure and count classifiers in Li's (2013) study of

Chinese classifiers respectively. Adopting Li's syntactic structures for measure and count classifiers, I propose the following two structures for prenominal duration and frequency phrases.

(4) The Prenominal Duration Phrase (5) The Prenominal Frequency Phrase
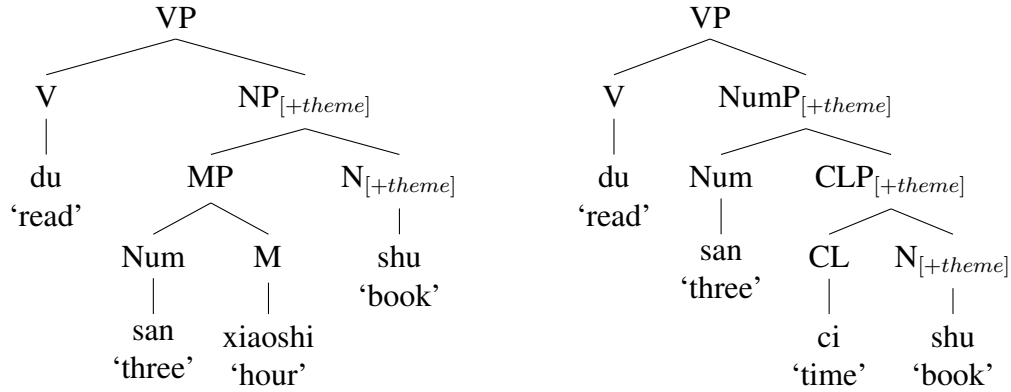
Based on this syntactic analysis, I further propose a direct-compositional semantic account of these two constructions that resolves the syntactic-semantics mismatch. I claim that the theme NP is shifted to an event description meaning, which can be measured out by the duration phrase, or counted by the frequency phrase.

(6) $[\![shu(\text{book})_{[+theme]}]\!]=\lambda e\exists x[theme'(e)(x) \wedge book'(x)]$
$[\![xiaoshi\ (\text{hour})]\!]=\lambda d\lambda e[duration'(e) = hour'(d)]$
$[\![ci\ (\text{time})]\!]=\lambda P_{<s,t>}\lambda d\lambda e[atom'(e) \wedge count'(e) = d \wedge P(e)]$

This paper is organized as follows: in §2, I discuss some previous syntactic analyses of duration and frequency phrases. Then, in §3, I present data to show in turn that these pronominal duration and frequency phrases are refined to the direct object position; unlike a double object construction, they indeed form a single constituent; and that these two are parallel to measure/count classifier constructions (Li 2013). The actual syntactic and semantic analyses come in §4, and the conclusion follows in §5.

2. Background

As is well documented in the Chinese linguistic literature, Mandarin duration phrases (DrP) and frequency phrases (FP) pattern almost identically in their syntax (Chao 1968; Huang et al. 2009): unlike other temporal adverbs, neither appears preverbally; both may appear postverbally either before or after the direct object as shown in (7) and in (8); when preposed before the verb, they trigger the verb-copying phenomenon as in (9). In all these three configurations, the meaning is roughly the same. Of these three constructions, this paper will mainly focus on sentences in (7), the prenominal DrP and FP. In this session, I will discuss a few major analyses of DrP and FP in turn: the analysis as an adjunct of V, the analysis as a complement of V, and the analysis as a numeral-classifier phrase.

(7) a.   Ta du le san xiaoshi (de)   shu.
          S/he read PFV three hour   (MOD) book

          'S/he read books (or a book) for three hours.'

     b.   Ta du le san ci   (de)*   shu.
          S/he read PFV three time (MOD)* book

          'S/he read books (or a book) three times.'

(8) a.   Ta du le na-ben shu san xiaoshi.
          S/he read PFV that-CL book three hour

          'S/he read that book for three hours.'

     b.   Ta du le na-ben shu san ci.
          S/he read PFV that-CL book three time

          'S/he read that book three times.'

(9) a.   Ta du shu du le san xiaoshi.
          S/he read book read PFV three hour

          'S/he read books (or a book) for three hours.'

     b.   Ta du shu du le san ci.
          S/he read book read PFV three time

          'S/he read books (or a book) for three times.'

2.1.   *Two previous analyses*

In previous studies, Mandarin duration and frequency phrases are predominantly analyzed in two ways: as an adjunct of V', or as a noun complement of the verb. Both of these theories analyze the three different word orders mentioned above of the duration phrase and the frequency phrase as derivable from a base-word order. However, neither of these two analyses can adequately explain the syntactic patterns of the prenominal duration and frequency phrases, because the prenominal duration and frequency phrases are essentially not derivable through movement from other word orders.

Under the V' adjunct analysis (Tang 1990; Huang et al. 2009), the duration phrase and the frequency phrase are adjuncts of V'. Because the complement of V can optionally move to the SPEC of VP, postverbally the duration and the frequency phrases can appear either before or after the object.

Although the V' adjunct analysis can indeed account for position variations postverbally, it is not refined enough to account for one difference between the object before and after the duration and frequency phrase. As Cheng et al. (1997:207) points out, 'The pattern [V object D/F] requires the object NP to be definite (or more precisely, a referential NP in Huang's terms)'. As shown in (11), when the direct object *shu* 'book' appears before the duration or the frequency phrase, it must be preceded by a demonstrative and a classifier.

(10)
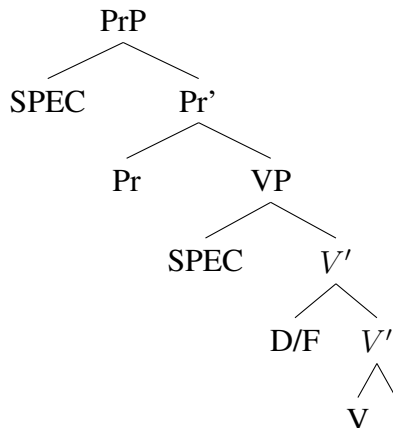
(11) a. * wo kan-le   (henduo) shu   liang ci/   liang-ge zhongtou.
          I   read-PFV many   book two   time/ two-CL   hour
          'I read (many books twice/for two hours.)'

      b.   wo kan-le   na-ben shu   liang ci/   liang-ge zhongtou.
          I   read-PFV that-CL book two   time/ two-CL   hour

          'I read that book twice/ for two hours.'

In order to account for this definiteness difference, Cheng et al. (1997) proposes that indefinite and definite NPs actually occupy different positions within the Mandarin VP structure: as shown in (12), an indefinite NP occupies the NP3 position and a definite NP occupies the NP2 position. Because under their analysis, the duration and frequency phrases are analyzed as noun complements of the verb, the duration and the frequency phrases actually compete for the NP3 argument slot with the direct object. Consequently, (11a) is ungrammatical precisely because the DrP or the FP cannot co-exist with the indefinite NP object in the NP3 position.

(12)
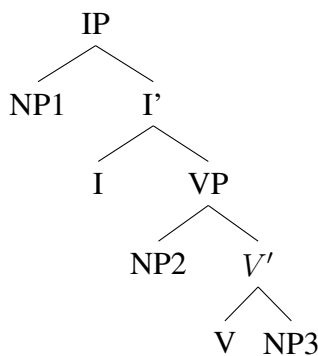
Cheng et al. (1997:209)

(13)   A definite/referential object NP occurs in the SPEC of VP position and an indefinite/non-referential object NP occurs within V' (as sister to V)

Cheng et al. (1997:209)

Besides accounting for the definiteness facts, the noun complement analysis has one more advantage over the V' adjunct analysis: it can also account for the verb copying sentences for

free. In the V' adjunct analysis, because the direct object and the duration/frequency phrase actually occupy different locations in the tree structure, verb copying is not necessary as a means to save a potentially ungrammatical structure, so that verb copying must be explained through some other mechanisms. In the noun complement analysis, because the the duration phrase or the frequency phrase competes for the same NP3 slot with the direct object syntactically, the direct object is forced to move upward dragging the verb along so as to make room, whenever NP3 is already occupied by the duration phrase or the frequency phrase.

Given these reasons above, I will adopt the noun complement analysis as a general framework for postnominal duration and frequency phrases and verb copying. However, as I will show in the next subsection, even the noun complement analysis encounters some problems for the prenominal duration and frequency phrases. This is where we need a third analysis, i.e. the 'numeral+classifier' analysis (Sybesma 1999; Liao 2014).

2.2. *Problems specific to the prenominal duration or frequency phrase*

As observed by Sybesma (1999), when a duration phrase or a frequency phrase appears before the object, they seem to behave like a 'numeral+classifier' phrase. This idea might seem unintuitive, but it actually has been hinted at by various authors such as Tang (1990) and Huang (1982) before Sybesma, and explicitly promoted by Liao (2014). In this subsection, I will focus on two important observations by Sybesma (1999): the referential/quantized constraint of the direct object and the parallelism to classifiers and massifiers. These properties would be quite hard to explain under the two analyses mentioned previously that do not assume that the direct object and the duration or frequency phrase form a single constituent.

First, Sybesma (1999) shows that the direct object in a prenominal duration or frequency phrase construction cannot be referential or quantized. As shown in the examples in (14) below, when the direct object 'book' follows the duration phrase, it cannot be preceded by a demonstrative *zhe* 'this' (not referential) or a numeral-classifier phrase such as *san ben* 'three CL' (not quantized). These facts can be easily explained if the duration phrase or the frequency phrase is analyzed as a 'numeral-classifier' construction, because syntactically the head noun cannot have another 'numeral-classifier' or a demonstrative, when the prenominal duration phrase and frequency phrase are already present as a 'numeral-classifier' phrase.

(14)   Duration of Time

    a.   Ta du le   san xiaoshi (de)   shu
            S/he read PFV three hour   (MOD) book
            'S/he read books (or a book) for three hours.'

    b. * Ta du le   san xiaoshi (de)   san-ben shu
            S/he read PFV three hour   (MOD) three-CL book
            Intended 'S/he read three books for three hours.'

    c. * Ta du le   san xiaoshi (de)   zhe-ben shu
            S/he read PFV three hour   (MOD) this-CL book
            Intended 'S/he read this book for three hours.'

Second, he crucially points out that the prenominal duration phrase and the prenominal frequency phrase pattern like mass-classifier and count-classifier respectively in respect of the

distribution of the genitive/modifier marker *de*. Although as discussed so far, the prenominal duration phrase and the prenominal frequency phrase are almost identical, they differ in one important place that only the duration phrase can have an optional genitive/modifier head *de* in-between (cf. Chao 1968; Jiang 2012). Sybesma (1999) cleverly links this distinction to the mass/count distinction in Chinese classifiers. This point will be discussed in more details in §3.4.

(15) a. wo du   le   san xiaoshi de   shu.
          I   read PFV three hour   MOD book.

          'I read books (or a book) for three hours.'

      b. wo du   le   san xiaoshi shu.
          I   read PFV three hour   book.

          'I read books (or a book) for three hours.'

(16) a. wo du   le   san ci   shu.
          I   read PFV three time book.

          'I read books (or a book) three times.'

      b. * wo du   le   san ci   de   shu.
          I   read PFV three time book.

          'I read books (or a book) three times.'

Although Sybesma (1999) has given a comprehensive description of the syntactic facts of the prenominal duration and frequency phrases, he has not given explicit structures for the prenominal duration and frequency phrases, nor has he provided any tests to prove that the prenominal duration and frequency phrases do indeed form a single constituent with the following object in the first place. Liao (2014) fills this gap by providing some indirect evidence from tonal sandhi in Taiwanese Chinese. But Liao (2014) also does not give a explicit syntactic structure for either. These authors do not spell out their analyses probably because the structure of a numeral classifier structure may seem to be too obvious, and that it is relatively hard to pin point the syntactic evidence that can tease these structures apart.

To fill this gap, in the next section, I introduce some new pieces of data to show some important properties of the prenominal duration or frequency phrase construction. I show that the prenominal duration and frequency phrases are restricted to only the direct object position, they do form a single constituent with the direct object with clear constituency test results, and that prenominal duration frequency phrases can be shown to have two distinct syntactic structures.

3. New data

In this section, I introduce a new set of data to show that the prenominal duration and frequency phrases do from a constituent with the direct object and that this constituent is indeed a 'numeral+classifier' phrase, but with two separate structures for the prenominal duration phrase and the prenominal frequency phrase.

Before discussing the constituency test results, I must first mention two caveats: head nouns with and without an inherent duration behave slightly differently and that these constructions are restricted to the direct object position only. These two points have not been discussed extensively in the literature but nevertheless affect the data patterns quite significantly.

Since previous studies surprisingly have not shown that these constructions are indeed constituents explicitly with the most basic constituency tests, I fill this gap by offering clear evidence from constituency tests. Crucially, these constituency tests are able to tease a base-generated and a derived constituent such as a double-object construction apart, so that we can rest assured that this constituent is a base-generated constituent.

Although this constituent has already been named 'numeral+classifier' construction in the previous studies (Sybesma 1999; Liao 2014), the actual syntactic structure has not been spelled out explicitly. With new tests from Li (2013), I show that the prenominal duration and the prenominal frequency phrases are parallel to counting classifiers and measuring classifiers in Li's (2013) study respectively. Consequently, because these two types of classifiers are argued to be structurally different in his study, I likewise claim that the prenominal duration and the prenominal frequency phrases have two different structures.

3.1.   *Different types of nouns and inherent duration*

Almost all types of indefinite nouns can be the direct object in the prenominal duration or frequency phrase construction. However, throughout the following subsections, the prenominal duration or frequency construction with certain types of nouns (such as 'movie', 'train ride') as the head noun seems to be more acceptable than that of other types of nouns (such as 'book'). And the reason is that these nouns differ in whether they have inherent duration and hence can be directly modified by a 'duration + de' phrase. For example, a noun such as 'movie' has a fixed duration, and so even outside of the specific constructions we are considering here, it can be modified by a 'duration +*de*' phrase. In contrast, a noun such as 'book' lacks an inherent duration, because in the real world a book is not associated with any specific amount of time. In-between these two extreme cases, a noun such as a 'train ride' can be loosely associated with a specific duration of time, because it takes about the same amount of time to get from one place to another, even though the actual amount of time taken is somewhat different each time.

(17)   a.   san   xiaoshi de   dianying
             three hour   MOD movie
             'movies of three hours' length'

       b.   * san   xiaoshi de   shu
             three hour   MOD book
             Intended 'books of three hours'

       c.   san   xiaoshi de   huoche
             three hour   MOD train
             'train rides of three hours'

Because these nouns differ in whether they can be directly modified by a 'duration + de' phrase already, in some cases sentences with nouns with an inherent length such as 'movie' seem to be more acceptable than their counterparts with nouns without an inherent length such as 'book'. It is therefore necessary to keep this in mind for our discussion of the new data below.

<center>*3.2. Restriction to direct object position*</center>

The construction of prenominal duration phrases or frequency phrases is restricted only to the direct object position. As the following example shows, this construction is not acceptable in the subject or the indirect object position.

(18)   (san   xiaoshi*) Mali gei-le   (san   xiaoshi*) Yuehan pingguo.
        (three hour*)   Mary give-PFV (three hour*)   John   apple.
        'Mary gave John apples for three hours.'

Interestingly, if the direct object is introduced by the *ba* marker (literally 'hold', a grammaticalized particle for introducing the direct object), the construction is not acceptable. This shows that the construction cannot be introduced by the direct object marker *ba*, and it must be syntactically a surface complement to V. This may suggest that the direct object in the 'DP/FrP + NP' form is no longer concrete semantically, because it is incompatible with *ba*.

(19)   a.   * mali ba san   xiaoshi shu   kan le.
              Mary BA three hour   book read PFV.

              Intended 'Mary read books (or a book) for three hours.'

        b.   * mali ba san   ci   shu   kan le.
              Mary BA three time book read PFV.

              Intended 'Mary read books (or a book) three times.'

The construction is equally infelicitous if passivized. This infelicity may be due to the fact that passives in Chinese are not as robust in English. Or it may show that this construction cannot go through A-movement. This also provides another piece of evidence that the construction is infelicitous in the surface subject position.

(20)   a.   * san   xiaoshi shu   bei   mali du le.
              three hour   book PASS Mary read PFV

              Intended 'Three hours' worth of books are read by May.'

        b.   * san   ci   shu   bei   mali du le.
              three time book PASS Mary read PFV

              Intended 'Three times of books are read by May.'

So from these facts above, we can see that the the prenominal duration and frequency phrases are highly restricted in their syntactic distribution: so far they are confined to the canonical direct object position. However, note that for nouns that have inherent duration of time, such as a movie or a train ride, the nouns themselves are already modifiable by the duration phrase. In these cases, the prenominal duration phrase with *de* is felicitous even in other positions. As (21) shows, the noun 'movie' with the modifying duration phrase is felicitous in the subject position and yet the same is not true for a ordinary noun as 'book'. Interestingly, the genitive/modifier marker *de* must be present in this case. This suggests that the prenominal duration phrase without *de* is confined to the postverbal prenominal position.

(21) a. * san xiaoshi de   shu hen youqu.
           three hour   MOD book very interesting
           Intended 'Books that last for three hours are very fun.'

      b.   san xiaoshi de   dianying hen youqu.
            three hour   MOD movie   very interesting
            Intended 'Movies that last for three hours are very fun.'

      c. * san   xiaoshi dianying hen youqu.
            three hour   movie   very interesting
   Intended 'Movies that last for three hours are very fun.'

Having shown that the prenominal duration or frequency phrases are confined to direct object position, I next turn to constituency tests to show that these constructions are base-generated single constituents.

3.3. *Constituency tests*

Besides in Liao's (2014) work, as far as I know, the prenominal duration or frequency phrases have not been shown explicitly as forming a single constituent with the following direct object. Whereas 'duration phrase+ *de*+ N' is mostly unanimously accepted as a single constituent, 'duration phrase/frequency phrase + N' has not been widely accepted as one constituent. Recall the two analyses of duration phrases above. Under either the V' adjunct analysis or the noun complement analysis, the preceding duration or frequency phrase does not form a single constituent with the following object. Instead, under the noun complement analysis, the prenominal duration phrase or frequency phrase resembles a double-object construction on the surface, because a double-object construction is of the structure 'V IO DO' in Mandarin as in (22). The two structures both look like there are two post-verbal elements after the verb, and these two elements do not form a constituent.

(22)   Mali gei-le   Yuehan yi-ge   pingguo.
        Mary give-PFV John   one-CL apple
        'Mary gave John an apple.'

As shown in the structure repeated below, NP2 and NP3 do not form a single base-generated constituent, but they do form a derived constituent once the verb moves out to I.

(12)

Cheng et al. (1997:209)
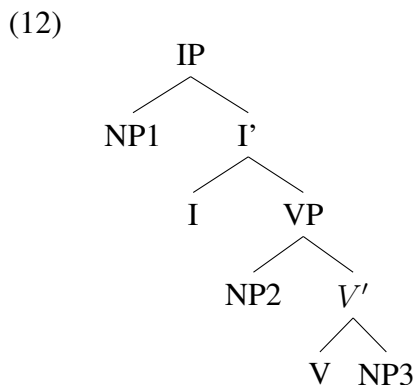
Indeed, as noted by Huang et al. (2009:83), as a derived constituent a double-object construction can pass certain constituency tests but not others. For example, as shown in (23), a double-object construction can be juxtaposed (Huang et al. 2009), even though it does not pass other constituency tests such as coordination, toplicalization or cleft.

(23)　? ta di-gei　gege　yi-hu　jiu, jiejie yi-pan cai.
　　　　 he pass-give brother one-CL wine, sister one-CL dish

　　　　 'He passed his brother a jug of wine and his sister a dish.'　　Huang et al. (2009:83)

If the noun complement analysis is correct, the prenominal duration or frequency phrase should also behave similarly to a double-object construction regarding the constituency tests. However, as the following examples show, unlike a double-object construction, the prenominal duration and frequency phrases pass almost all constituency tests such as coordination, topicalization, and cleft. As each of the following examples show, only the 'duration+N' phrase but not the double-object construction passes all these tests.

(24)　Coordination. (*he* conjoins NP and *erqie* conjoins clauses or VPs)

　　　a.　　Mali kan le　san　xiaoshi shu　he/erqie* yi　xiaoshi dianshi.
　　　　　　Mary watch PFV three hour　　book and　　one hour　TV

　　　　　　'Mary read for three hours and watched TV for three hours.'　　　duration+N

　　　b.　* Mali gei le　Bide yi-ge　juzi　he*/erqie* Yuehan yi-ge　pingguo.
　　　　　　Mary give PFV Peter one-CL orange and　　　John　one-CL apple

　　　　　　Intended 'Mary gave Peter an orange and John an apple.'　　double objects

(25)　Topicalization.

　　　a.　　san　xiaoshi shu　wo dique　kan-le.
　　　　　　three hour　book I　indeed watch-PFV.

　　　　　　'I indeed read books for *three hours*.'　　　　　duration+N

　　　b.　* Yuehan yi-ge　pingguo wo dique　gei-le.
　　　　　　John　one-CL apple　I　indeed give-PFV.

　　　　　　Intended 'I indeed gave an apple to John.'　　　double objects

(26)　Cleft

　　　a.　　wo kan　le　de　shi san　xiaoshi shu.
　　　　　　I　watch PFV REL COP three hour　book.

　　　　　　Roughly 'What I did was reading books for three hours.'　　　duration+N

　　　b.　* wo gei le　de　shi Yuehan yi-ge　pingguo.
　　　　　　I　give PFV REL COP John　one-CL apple

　　　　　　Intended 'What I did was giving John an apple.'　　　double objects

So these tests provide strong evidence that the prenominal duration phrase or frequency phrase forms a single constituent with the following direct object. And this constituent is not a derived constituent like the two objects of a double-object construction. The 'DrP/FP phrase + N' construction is a single constituent without any verb movement outward.

To digress slightly, these constituency tests actually also bear on the syntactic status of postverbal duration and frequency phrases. Interestingly, unlike their preverbal counterparts, the postverbal duration and frequency phrases cannot pass the above constituency tests. These

results show that like a double-object construction, postverbal duration and frequency phrases do not form a single constituent, lending support to Cheng et al.'s (1997) analysis of postverbal frequency and duration phrases as a double-object construction. However, if the preverbal and postverbal duration and frequency phrases are indeed related by movement, they should both behave similarly regarding the constituency tests, assuming that both share the same underlying structure. This is contrary to what we find here.

(27) * Mali kan le   zhe-ben shu   san-xiaoshi he/erqie dianshi yi   xiaoshi .
       Mary look PFV this-CL book three-hour and   TV   one hour

      Intended 'Mary read this book for three hours and watched TV for an hour.'

(28) * zhe-ben shu   san-xiaoshi wo dique   kan-le.
       this-CL book three-hour I   indeed look-PFV

      Intended 'I indeed read this book for *three hours*.'

(29) * wo kan le   de shi zhe-ben shu   san   xiaoshi .
       I   look PFV REL COP this-CL book three hour

      Intended 'What I did was reading this book for three hours.'

Thus, because the prenominal duration and frequency phrases are probably not related to the postnominal ones by movement, they call for a different analysis from Cheng et al.'s (1997) double-object analysis.

3.4.   Parallelism to Chinese measure phrases

In the last section, it has been established that prenominal duration and frequency phrases indeed form a single constituent with the direct object. According to Sybesma (1999) and Liao (2014), this constituent is a 'numeral + classifier' construction, because the prenominal duration and frequency phrases are structurally akin to two different kinds of classifier phrases.

3.4.1.   Two types of classifiers

A classifier phrase in Mandarin is of the structure 'numeral + classifier + head noun'. A classifier is mandatory between the numeral and the head noun, even if the head noun is countable. For example, as shown in (30), the classifier *ge* must be present between 'three' and 'person', otherwise the sentence is ungrammatical. This type of classifiers for countable nouns are considered to be the prototypical classifiers, and it is called individual classifiers by Li (2013). Another type of classifiers, called 'measure classifiers' by Li (2013), are measure words that express a standard unit, such as 'pound' or 'meter'. In other words, a measure phrase 'three pounds of beef' and a numeral-classifier construction 'three books' in Mandarin share the same structure 'numeral+classifier+head noun'. According to Cheng & Sybesma (1998), these two types of classifiers can be differentiated by whether *de* the genitive/modifier marker is present. Of these two, only measure classifiers can have an optional *de* intervening. Because these two types of classifiers modify count and mass nouns respectively, Cheng & Sybesma (1998) propose that whether *de* can appear in a classifier phrase depends crucially on this mass/count distinction of the head noun.

(30) Individual Classifiers

    a.   san (ge)* ren
            three (CL)* person
            'three people'

    b. * san ge de   ren
            three CL MOD person
            'three people'

(31) Measure Classifiers

    a.   san bang niurou
            three pound beef
            'three pounds of beef'

    b.   san bang de   niurou
            three pound MOD beef
            'three pounds of beef'

And yet, it has been questioned whether the distribution of *de* really reflects the count/mass distinction of classifiers by several authors (Hsieh 2008; Jiang 2009; Li 2013, among others). Li (2013), for example, proposes that *de* actually distinguishes the counting and the measuring meaning of classifiers instead, and some classifiers may have either the counting or the measuring reading depending on the context.

Intuitively speaking, the counting reading denotes physically concrete discrete objects, where as the measuring reading denotes an intangible amount. Under certain circumstances,, classifiers that are usually only used in the counting reading may be coerced into a measuring reading. Consider the following examples from Li (2013). A container measure word 'bottle' is ambiguous between a counting and a measuring reading. For example,in (32a), the classifier phrase *liang ping jiu* 'two bottles of wine' have a counting reading , because the two bottles are physically tangible objects; whereas in (32b), it has a measuring reading, because the 'two bottles of wine' refers to the amount of wine instead of the physical tangible objects of two bottles of wine.

(32) a.   wo ling le liang ping   jiu, zuo shou yi ping, you shou yi ping.
            I   lift PFV two CL_{bottle} wine left hand one CL_{bottle} right hand one CL_{bottle}
            'I carried two bottles of wine, one in the left hand and the other in the right hand.'

    b.   ta-de jiuliang   shi liang ping   hongjiu.
            his   drinking-capacity be two CL_{bottle} red-wine
   'His drinking-capacity is two bottles of red wine.'   Li (2013:135)

So for Li (2013), *de* can be present in a classifier phrase just in case the classifier is interpreted with a measuring reading. As the following two examples show, *de* is felicitous in the context where only the measuring reading is available. It is infelicitous in a context about opening three bottles of wine, but it is felicitous in a context where the 'three bottles' refers to an amount.

(33) a.   fuwuyuan kai le san ping   (#de) jiu.
            waiter   open PFV three CL_{bottle} DE   wine
            'The waiter opened three bottles of wine.'   [Counting]

    b.   zhe ge peng keyi zhuan san ping   (de) jiu.
            this CL basin can contain three CL_{bottle} DE wine
            'This basic can contain three bottles of wine.'   [Measure]
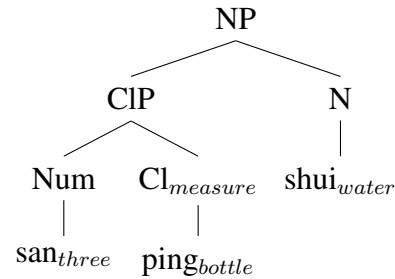
                                              Li (2013:139)

This piece of evidence, as Li (2013) argues, shows that the numeral and the classifier form a constituent in the measuring reading, but not in the counting reading. He proposes the following two structures for a counting reading and a measuring reading.
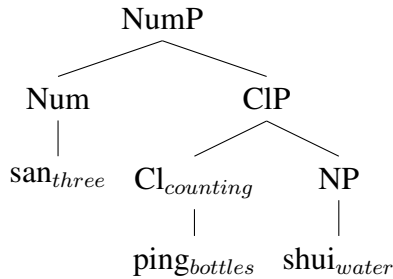
(34)   Counting   (35)   Measuring

Besides the distribution facts of *de*, Li (2013) provides another test with *duo* to show that the two types of classifiers are indeed syntactically distinct. He observes that the additive modifier *duo* 'more' only has a measuring reading when added between the classifier and the noun. For most numeral-classifier constructions, *duo* can appear before or after the classifier.

(36)   a.   shi duo   feng xin
            ten more CL   letter
            'more than ten letters'

       b.   shi chi   duo bu
            ten CL$_{inch}$ more cloth
            'more than ten inches of cloth'   adapted from (Li 2013:137)

When *duo* appears before the classifier, the numeral-classifier construction can either have a counting or a measuring reading.

(37)   a.   ta ling le   shi duo ping   hong jiu.
            he carry PFV ten more CL$_{bottle}$ red   wine
            'He carried more than ten bottles of red wine.'   [counting]

       b.   ta zhi shao neng hexia shi duo   ping   hong jiu.
            he at   least can   drink ten more CL$_{bottle}$ red   wine
            'He can at least drink more than ten bottles of red wine.'   [measure]

                                                                        Li (2013:1387)

However, when *duo* appears after the classifier, the numeral-classifier construction can only have a measuring reading. The following contrast in (38) and (39) shows that when *duo* appears after the classifier, the classifier construction cannot denote concrete objects, which are the actual bottles of wine, making (39) infelicitous in this case. Instead, the classifier phrase with *duo* after itself can only denote some abstract quantity of wine with bottle as a counting unit so that (38) is felicitous.

(38)   ta zonggong he   le   you san ping   duo hong jiu.
            he altogether drink PFV have three CL$_{bottle}$ more red   wine
            'He drank more than three bottles of red wine.'   [measuring]

(39) # ta ling le san ping duo hong jiu.
        he carry PFV ten CL*bottle* more red   wine
        Intended: 'He carried more than ten bottles of red wine.'   [counting]

<div align="right">adapted from Li (2013:138)</div>

Li (2013) explains these distributional facts about *duo* as follows: *duo* can either attach to the numeral alone, making the number not exact, or it can attach to the 'numeral classifier' constituent in the measuring reading, making the amount not exact. It does not make sense to attach *duo* to the classifier alone, because the classifier cannot be made imprecise semantically speaking.

So from the distribution facts of *de* and *duo*, we can conclude that indeed classifiers with two different meanings, counting and measuring, have two different syntactic structures. Likewise, in the following subsection, the prenominal duration and frequency phrases will be shown to behave exactly parallel to these two types of classifiers respectively with respect to the *de* test and the *duo* test.

3.4.2. *Parallelism to two types of classifier phrases*

According Sybesma (1999), like the two types of classifier phrases mentioned above, the prenominal duration and frequency phrases behave differently with respect to whether *de* can be present. Only the prenominal duration phrase can have *de* intervening. From the parallels between (40) and (41), and between (42) and (43), it is clear that the prenominal duration phrase is like a measuring classifier phrase, whereas the prenominal frequency phrase is like a counting classifier phrase.

(40) Measuring Classifier Phrase

    a.   san gongjin lizhi
           three kilo   lychee
           'three kilos of lychees'

    b.   san gongjin de   lizhi
           three kilo   MOD lychee
           'three kilos of lychees'
           Or '(some) three-kilo lychee'

(41) Duration Phrase +N

    a.   kan san xiaoshi dianshi
           watch three hour   TV
           'watch TV for three hours'

    b.   san   xiaoshi de dianshi
           three-hour MOD   TV
           'watch TV for three hours'
           (lacks a kind reading)

(42) Counting Classifier Phrase

    a.   san ge ren
           three CL people
           'three people'

    b.   * san ge de   ren
           three CL MOD people
           Intended 'three people'

(43) Frequency Phrase +N

    a.   kan san ci dianshi
           watch three time TV
           'watch TV three times'

    b.   * san ci de   dianshi
           three time MOD TV
           Intended 'watch TV three times'

Although Sybesma (1999) makes this brilliant association, he actually does not spell out how the prenominal duration phrase and the prenominal frequency phrase differ in their syntactic structures. As he assumes that these two types of classifiers differ in their selectional restrictions

for an optional *de*, he claims that the prenominal duration and frequency phrases only differ in their selection restrictions but not in their structures. However, from the discussions of Li's (2013) arguments in the last subsection, counting and measuring classifier phrases clearly have two different syntactic structures. As the prenominal duration and frequency phrases are parallel to these two different types of classifiers with respect of the distribution of *de*, they too should have two different syntactic structures like in Li's proposal.

The prenominal duration phrase and frequency phrases are also parallel to measuring and counting classifier phrase respectively with respect to the distribution of *duo* 'much'. Recall that *duo* 'much' can be inserted before a classifier freely, and after a classifier only when the classifier phrase has a measuring reading. Likewise, both the prenominal duration and the prenominal frequency phrases can have *duo* inserted before the classifier, while only the prenominal duration phrase can have *duo* following the duration of time word such as 'hour'. As the following examples show, *duo* can appear before both the duration measure word *xiaoshi* 'hour' and the frequency count word *ci* 'time', and yet it can only appear after *xiaoshi* 'hour' but not *ci* 'time'.

(44) a. wo du-le   shi duo   xiaoshi (de)   shu.
          I   read-PFV ten much hour   (MOD) book.
          'I read books (or a book) for about ten hours.'

      b. wo du-le   shi duo   ci   shu.
          I   read-PFV ten much time book.
          'I read books (or a book) about ten times.'

(45) a. wo du-le   shi xiaoshi duo   (de)   shu.
          I   read-PFV ten hour   much (MOD) book.
          'I read books (or a book) for about ten hours.'

      b. * wo du-le   shi ci   duo   shu.
            I   read-PFV ten time much book.
   'I read books (or a book) about ten times.'

Beyond these syntactic parallelisms, there are good semantic reasons why the prenominal duration and frequency phrases may be parallel to classifier phrases. Just as numeral classifier phrases count or measure physical entities, frequency phrases and duration phrases count or measure events. A frequency phrase counts the number of occurrences of an event, and therefore the word *ci* 'time' can be considered to be the counting classifier of an event; whereas a duration phrase measures the duration of an event, and consequently duration measure words such as *xiaoshi* 'hour' can be interpreted as a measuring classifier for an event. Consequently, the head noun in the prenominal duration and frequency phrases should perhaps be considered as denoting an event, rather than simply as physical objects.

For my proposal below in the next section, I adopt and modify Li's (2013) analysis of numeral classifiers for the prenominal duration and frequency phrases. Based on this syntactic analysis, I propose a compositional semantic account that the head NP is shifted into an event type meaning, which resolves the syntactic and semantic mismatch.

4. Proposal

Adopting Li's (2013) syntactic structures, I propose the following two structures for the prenominal duration phrase and the prenominal frequency phrase. These two structures are exactly the same as the structures for the measure reading and the count reading in Li (2013), except that the direct object NP has a special [+theme] feature. This feature is assigned by the verb to its direct object to make sure that the construction is only licensed in the direct object position, and that the direct object has an event interpretation. Crucially, these two structures differ in that the numeral and the classifier form a constituent in the duration phrase, but not in the frequency phrase. As shown in (46) and (47), a duration measure word, such as *xiaoshi* 'hour', first forms an measure phrase (MP) constituent with the numeral, and then combines with the head noun with a $[+theme]$ feature, whereas a frequency classifier, such as *ci* 'time', first combines with the head noun to form a classifier phrase (CLP) and then combines with the numeral to form a numeral phrase.[1]

(46)   The Prenominal Duration Phrase   (47)   The Prenominal Frequency Phrase
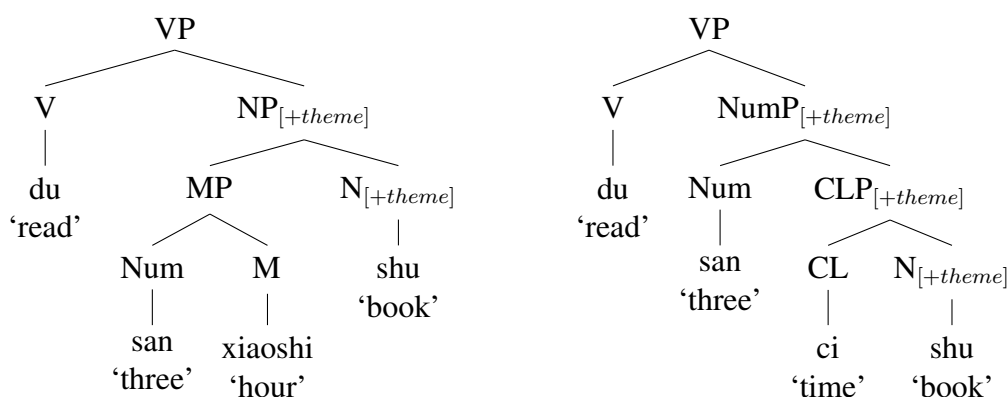
With these two new structures, we can now easily account for the data discussed in the previous sections: the restriction that the direct object cannot be quantized, the restriction of the prenominal duration and frequency phrases to the direct object position, the distribution of *de* the genitive/modifier marker, and the distribution of *duo* 'much'.

To begin with, in my analysis, the head noun is an N not a NP. This fact explains why, as discussed in the background section, the head noun cannot be quantized and referential. The head noun cannot have another 'numeral-classifier' in front to be quantized because there is already another numeral- classifier phrase for the duration or frequency in front. Neither can it be preceded by a demonstrative because a demonstrative selects for an NP or CLP. Recall the following example from the background section.

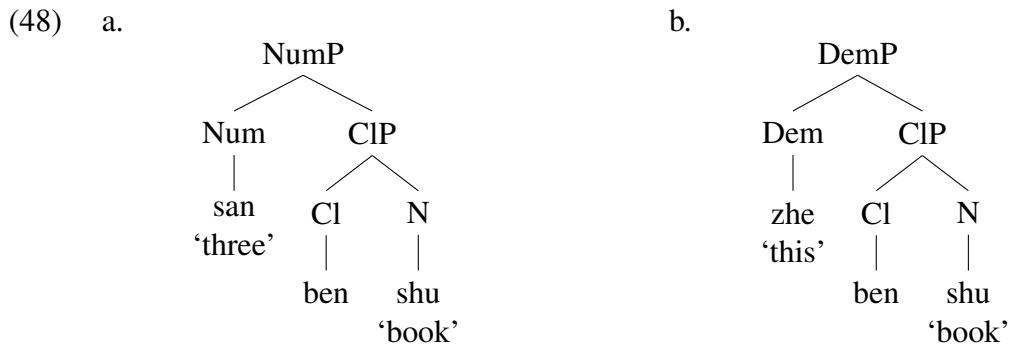(14)   Duration of Time

    a.   Ta du le san xiaoshi (de) shu
            S/he read PFV three hour   (MOD) book

        'S/he read books (or a book) for three hours'

---

[1] I have chosen to call $CL_{measure}$ and CL_{count} as M and CL respectively instead, because I think in Li's (2013) proposal, a ClP can be either 'Num + Cl' or 'CL + N' depending on the structure, which may be confusing.
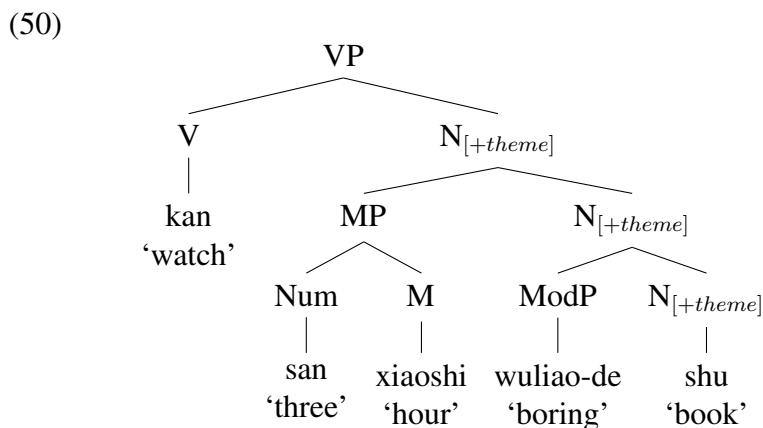
b. * Ta du   le   san   xiaoshi (de)   san-ben shu
  S/he read PFV three hour   (MOD) three-CL book
  Intended 'S/he read three books for three hours'

c. * Ta du   le   san   xiaoshi (de)   zhe-ben shu
  S/he read PFV three hour   (MOD) this-CL book
  Intended 'S/he read this book for three hours'

Both (14a) and (14b) are ungrammatical because after the duration phrase the noun is not a bare noun. Both *san ben shu* 'three books', and *zhe ben shu* 'this book' have projections larger than an N. As shown in the following example, the former has a NumP projection, and the latter has a DemP projection. In the new structures I have proposed, both the duration phrase *san xiaoshi* 'three hour' and the frequency classifier *ci* 'time' selects for an N$_{[+theme]}$ and consequently the syntactic derivation would crash, given that a quantized or a referential NP is a wrong input for the duration phrase and the frequency classifier.

(48)   a.

```
            NumP
           /    \
        Num      ClP
         |      /   \
        san   Cl     N
      'three'  |     |
              ben   shu
                   'book'
```

b.

```
            DemP
           /    \
        Dem      ClP
         |      /   \
        zhe   Cl     N
      'this'  |     |
              ben   shu
                   'book'
```
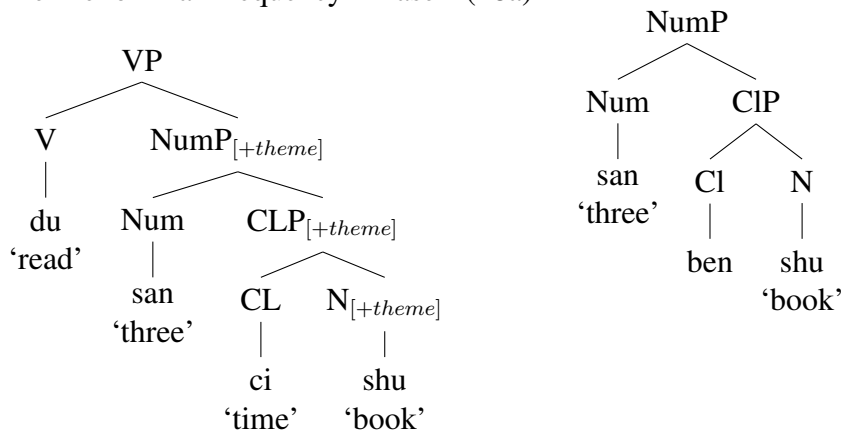
However, assuming that adjectives and relative clauses can modify bare nouns in Mandarin, my proposal above does not rule out the possibility of adjectives or relative clauses in front of the direct object. Indeed, this prediction is borne out. As (49) shows, an adjective such as *wuliao-de* 'boring' can intervene between the duration phrase and the head noun.

(49)   wo kan le   san   xiaoshi wuliao-de   dianying.
  I   read PFV three hour   boring-MOD movie
  'I watched boring movies (or a movie) for three hours.'

(50)

```
                    VP
                  /    \
                 V      N[+theme]
                 |      /      \
               kan    MP        N[+theme]
             'watch'  /  \      /      \
                   Num   M   ModP      N[+theme]
                    |    |     |          |
                   san xiaoshi wuliao-de  shu
                 'three' 'hour' 'boring' 'book'
```

Second, my analysis claims that the head noun has a $[+theme]$ feature. This explains why the prenominal duration and frequency phrases cannot appear in other syntactic positions except for the direct object position: other argument positions are not assigned the $[+theme]$ feature, it is syntactically as well as semantically infelicitous to measure an ordinary noun without inherent duration with time measurement phrases. Syntactically, an ordinary noun without a $[+theme]$ feature would not take *ci* 'time' as its classifier. For example, the classifier selected for *shu* 'book' should be *ben* not *ci*. Compare again (47) and (48a), two syntactic structures for a numeral-classifier structure with a $[+theme]$ feature and without a $[+theme]$ feature: the classifier for *shu* 'book' without a $[+theme]$ feature is *ben*; whereas with a $[+theme]$, it is *ci*.

(47) The Prenominal Frequency Phrase   (48a)

```
                VP                              NumP
              /    \                          /      \
            V       NumP[+theme]            Num       ClP
            |      /      \                  |        /   \
           du    Num       CLP[+theme]      san      Cl    N
         'read'   |       /      \         'three'    |     |
                 san    CL        N[+theme]          ben   shu
               'three'  |          |                     'book'
                        ci         shu
                      'time'      'book'
```

Semantically, a noun without an inherent duration of time such as 'book' cannot be felicitously modified by a duration phrase. However, at the direct object position, this $[+theme]$ feature gives the noun an event type of meaning, so that the noun is actually standing for an event with that noun as a participant. Because N[+theme] has an event type meaning, the classifier that modifies this N[+theme] is automatically specified as count classifiers of events, namely *ci* 'time' or other classifiers that expresses frequency. Likewise, N[+theme] with an event meaning can be measured by duration phrases such as *san xiaoshi* 'three hour'.

Third, according my proposal, only the duration phrase forms a constituent of its own. This explains the distribution of *de* and *duo* just as Li (2013) does for the two different types of classifiers.

For the distribution of *de*, recall the following examples from the previous section. *De* can only be inserted between a duration phrase and the head noun, but not between a frequency phrase and the head noun.

(41) Duration Phrase +N

    a.   kan   san xiaoshi dianshi
           watch three hour   TV
           'watch TV for three hours'

    b.   san xiaoshi de   dianshi
           three hour   MOD TV.
           'watch TV for three hours'
           (lacks a kind reading)

(43) Frequency Phrase +N

    a.   kan   san ci   diansh
           watch three time TV
           'watch TV three times'

    b. * san ci   de   dianshi
           three times MOD TV
           Intended 'watch TV three times'

Since the duration phrase has the syntax of MP as a constituent, *de* the modifier head can take DrP has a complement. In contrast, because the classifier and the numeral in the structure for the prenominal frequency phrase construction does not form a constituent, *de* is not allowed after the frequency phrase.
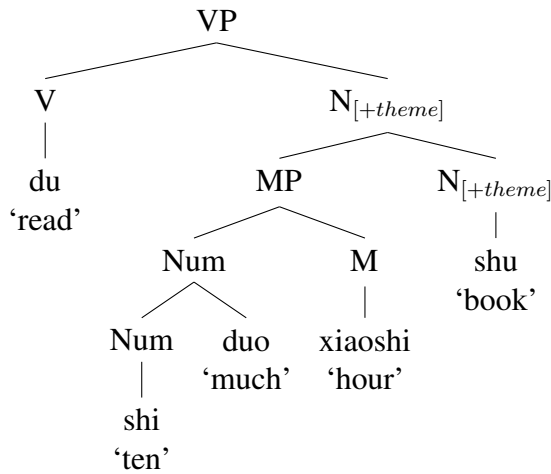
(51)
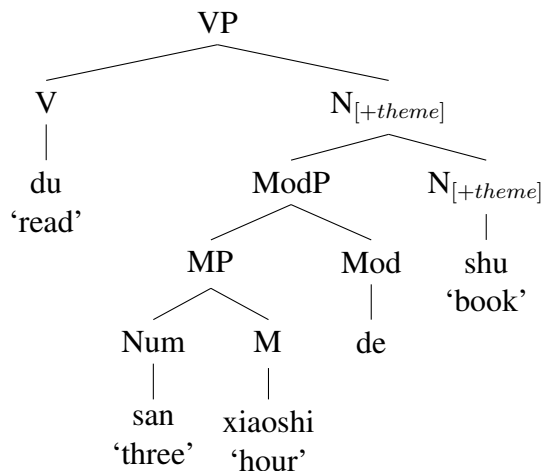
Similarly, the position of *duo* can be explained in the same fashion as follows. Recall the following examples (44) and (45) from the previous section. Before after the numeral, *duo* can appear both in the prenominal duration and the prenominal frequency phrase; whereas after the duration or frequency word, *duo* can only appear in the prenominal duration phrase.
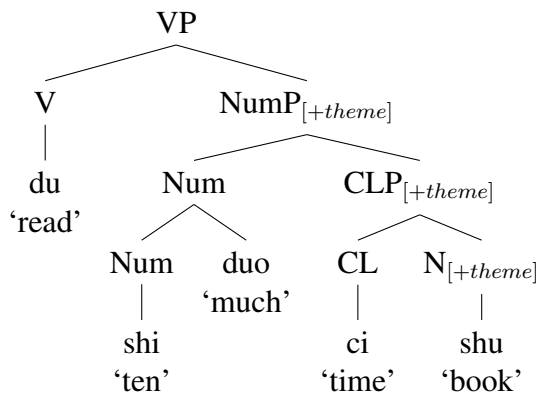
(44)   a.   wo du-le   shi duo   xiaoshi (de)   shu.
           I   read-PFV ten much hour   (MOD) book
           'I read books (or a book) for about ten hours.'

       b.   wo du-le   shi duo   ci   shu
           I   read-PFV ten much time book
           'I read books (or a book) about ten times.'

(45)   a.   wo du-le   shi xiaoshi duo   (de)   shu.
           I   read-PFV ten hour   much (MOD) book
           'I read books (or a book) for about ten hours.'

       b. * wo du-le   shi ci   duo shu.
           I   read-PFV ten time much book
           'I read books (or a book) about ten times.'

According to Li (2013), *duo* can appear immediately after the numeral because it can attach to the numeral directly. Since the numeral by itself is a constituent in both the prenominal duration and the prenominal frequency phrases, it is felicitous after the numeral in both constructions.
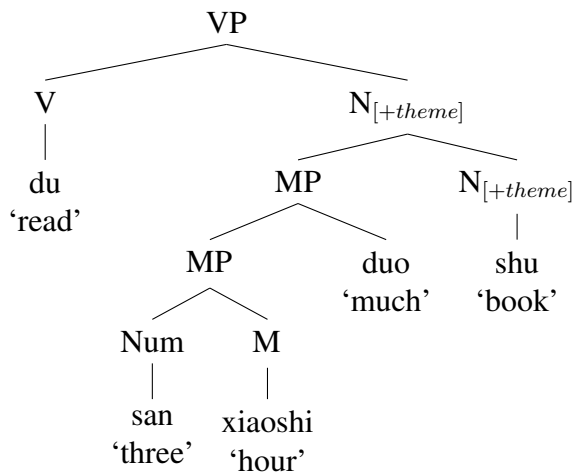
(52)

```
              VP
         ┌────┴──────────┐
         V              N[+theme]
         │          ┌──────┴───────┐
         du        MP            N[+theme]
       'read'   ┌───┴────┐          │
              Num        M         shu
          ┌────┴───┐     │        'book'
         Num      duo  xiaoshi
          │      'much'  'hour'
         shi
        'ten'
```

(53)

```
              VP
         ┌────┴──────────┐
         V            NumP[+theme]
         │        ┌──────┴────────┐
         du      Num            CLP[+theme]
       'read' ┌───┴───┐      ┌─────┴─────┐
            Num      duo     CL       N[+theme]
             │      'much'    │          │
            shi             ci          shu
           'ten'           'time'      'book'
```

Besides, *duo* can also attach to a 'numeral+classifier' constituent (Li 2013). Because the numeral and classifier form a constituent in the duration phrase but not in the frequency phrase, only after the duration phrase can *duo* be attached as a measure phrase modifier.

(54)

```
              VP
         ┌────┴──────────┐
         V              N[+theme]
         │          ┌──────┴───────┐
         du        MP            N[+theme]
       'read'   ┌───┴─────┐         │
              MP         duo       shu
          ┌────┴───┐   'much'     'book'
         Num        M
          │         │
         san      xiaoshi
        'three'    'hour'
```

My proposal is thus an improvement over previous analyses by explicitly spelling out two different structures for the prenominal duration phrase and the prenominal frequency phrase that covers much more data, such as the construction's restriction to the direct object position and

the distributional facts of *de* and *duo*. In the following subsection, I present my compositional-semantic account of the prenominal duration or frequency phrases based on this syntactic analysis.

4.1. Semantic composition

In the previous section, I have shown the syntax of 'durational phrase + (de)+DO' is parallel to measuring classifiers (massifiers), while 'frequency phrase + DO' is parallel to counting classifiers in Mandarin (cf. Sybesma 1999; Li 2013). I propose further that this syntactic proposal provides direct syntactic evidence in support of Kennedy's (2012) semantic analysis that incremental theme is an event measurement function, because it shows that an event variable can be associated with the direct object in Mandarin indeed.

Because in the prenominal duration or frequency phrase construction, the event measurement phrases are directly modifying the direct object, it is evident that the semantics of the direct object is an event description under a direct-composition analysis. I propose that a noun with a [+theme] feature is given a shifted meaning of an event description. As shown in (55), *shu* 'book' with a [+theme] feature is shifted to an event description that describes an event that has some book as the theme.

(55) $[\![shu(\text{book})_{[+theme]}]\!] = \lambda e \exists x [theme'(e)(x) \wedge book'(x)]$

Because the semantics of the theme in this construction is an event description, it can be measured out by a duration phrase or counted by a frequency phrase. Parallel to the semantics of the measuring classifiers and the counting classifiers in Mandarin, I propose that the semantics of a duration noun such as *xiaoshi* 'hour' is the measure function of the duration of an event by the unit of hour as in (56).

(56) $[\![xiaoshi(\text{hour})]\!] = \lambda d \lambda e [duration'(e) = hour'(d)]$

So the composition of *du san-xiaoshi shu* 'read three hours' book' is as follows: *xiaoshi* 'hour' first composes with the numeral three to render an event description that describes a three-hour event, and then this description composes with *shu* 'book' with the shifted event description meaning, and this renders a new event description that describes an event that lasts for three hours, and with some book as the theme. Finally, this new event description composes with *du* 'read' and renders an event description of a book-reading event that lasts for three hours.

(57) Semantic Composition for the Prenominal duration Phrase

$$\lambda e[read'(e) \wedge duration'(e) = hour'(3) \wedge \exists x[theme'(e)(x) \wedge book'(x)]]$$

$\lambda e.[read'(e)]$ $\lambda e[duration'(e) = hour'(3) \wedge \exists x[theme'(e)(x) \wedge book'(x)]]$

$\lambda e[duration'(e) = hour'(3)]$ $\lambda e \exists x[theme'(e)(x) \wedge book'(x)]$

3 $\lambda d \lambda e[duration'(e) = hour'(d)]$

Similarly, I propose that the frequency temporal noun *ci* 'time' is an event classifier that first atomizes an event description into countable instances, and then takes in a degree argument that counts the number of the occurrences of the event.

(58) $[\![ci \text{ (time)}]\!]=\lambda P_{<s,t>}\lambda d\lambda e[atom'(e) \wedge count'(e) = d \wedge P(e)]$

So the semantic composition of 'read three times' book' in Mandarin works as follows. The time frequency word *ci* 'times', first takes in the event description and atomizes the event description into countable individual events, and then the new event description that is countable directly composes with the numeral three to yield three instances of book-related events. Finally, after composing with *du* 'read', the meaning of 'read three times' book' is an event description that describes three individual events of book reading.

(59) Semantic Composition for the Prenominal Frequency Phrase

$\lambda e[read'(e) \wedge atom'(e) \wedge count'(e) = 3 \wedge \exists x[theme'(e)(x) \wedge book'(x)]]$

$\lambda e.[read'(e)] \quad \lambda e[atom'(e) \wedge count'(e) = 3 \wedge \exists x[theme'(e)(x) \wedge book'(x)]]$

$\mathbf{3} \qquad \lambda d\lambda e[atom'(e) \wedge count'(e) = d \wedge \exists x[theme'(e)(x) \wedge book'(x)]]$

$\lambda P_{<s,t>}\lambda d\lambda e[atom'(e) \wedge count'(e) = d \wedge P(e)] \quad \lambda e\exists x[theme'(e)(x) \wedge book'(x)]$

So my semantic account gives evidence in support of Kennedy's (2012) analysis of incremental change, that the theme argument is associated with an event variable in Mandarin. The theme argument describes a type of events that has that noun in the direct object position as the theme. However, unlike Kennedy's (2012) account, the event measurement function in Mandarin is not encoded in the noun itself, but in the duration and frequency words, the event measurement phrases themselves. Because the duration and the frequency phrases would not be able to directly modify the direct object, if the incremental theme in Mandarin directly encodes the measurement of change function. This shows an interesting cross-linguistic semantic variation in the domain of events in Mandarin. It seems to suggest that events in a language are also structured similarly to the noun phrases in that a language. Because in my proposal, the event description is like a mass noun, which cannot be directly counted, but can be measured. In order to be counted, the event classifier *ci* must be used. This reflects the analysis of Mandarin nouns in Chierchia (1998) that Mandarin nouns denote kind and must be atomized before it can be counted. In contrast, English nouns can be counted directly by numerals. Likewise, the measurement of change function lies in Mandarin event modifiers, whereas it lies in the incremental theme itself in English.

4.2. *Some Further Issues*

There remain some further issues to be explored with my current proposal. Due to the limited space, I address some issues briefly in this subsection.

The first issue is how general this syntactic account with the [+theme] feature can be. In particular, is it the case that every direct object can have this [+theme] feature. My answer is that the [+theme] feature is restricted and can be only assigned by durative verbs. Therefore, only activities, temporary states and accomplishments can assign such as potential [+theme] feature to its direct object. In this sense, I am using the [+theme] feature in a broader sense that the direct object provides something which the event can be measured out against. The purpose of this [+theme] feature is to mark the direct object as having a special meaning. One anonymous reviewer provides a valid point that the [+theme] may not be that special, and I agree that other mechanism might be able to produce the same results. The crucial point, however, is

that the indefinite NP modified by the duration or the frequency phrase must have an event-type of meaning. [2]

Another issue regards how this analysis of the semantics of the duration and the frequency phrase can extend to other word order variations. While my account of the direct object having a shifted event meaning is restricted to this particular construction, the semantics of the duration and the frequency phrases themselves are regular event descriptions, and the semantics of these phrases remain the same regardless of their actual positions. Given Cheng et al.'s (1997) syntactic structure, when the duration phrase or the frequency phrase occupies the complement of *V* position, semantically speaking the composition is now between two event descriptions, whereas the direct object is introduced later in another position.

5.   *Conclusion*

This paper gives a new syntactic-semantic interface account of the prenominal duration and frequency phrases in Mandarin. I have argued that the prenominal duration phrase and the prenominal frequency phrase form a constituent with the following theme NP, which provides direct syntactic evidence that under certain conditions the direct object in Mandarin can have an event type of meaning, because it can be directly modified by event measurement phrases. Therefore, this paper supports Kennedy's (2012) semantic analysis of incremental change. Furthermore, this paper shows that Mandarin and English encode the measurement of change function in different places: while English encodes the function in the theme NP itself, Mandarin encodes it in the event classifiers and measurement phrases themselves. This also suggests that events and noun phrases might be structured in a parallel fashion in a language, inviting potential future research into this topic.

Acknowledgements

Abbreviations

BA	Chinese Direct Object Marker.
CL	Classifier
COP	Copular
MOD	Modifier Head.
PASS	Passive
PFV	Perfective
REL	Relativizer

---

[2]Many thanks for the anonymous reviewer who provided very insightful and helpful feedback.

References

Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. University of California Press.

Cheng, L., C.-T. J. Huang, Y.-H. A. Li & C.-C. J. Tang (1997). Causative compounds across Chinese dialects: a study of Cantonese, Mandarin and Taiwanese. *Chinese Languages and Linguistics* 4, pp. 199–224.

Cheng, L. L.-S. & R. Sybesma (1998). Yi-wan tang, yi-ge tang: Classifiers and massifiers. *Tsing Hua Journal of Chinese Studies* 28:3, pp. 385–412, URL http://arca.unive.it/bitstream/10278/3195/1/Classifiers&Massifiers(ChengSybesma).pdf.

Chierchia, G. (1998). Reference to kinds across language. *Natural Language Semantics* 6:4, pp. 339–405, URL http://link.springer.com/article/10.1023/A:1008324218506.

Hsieh, M.-L. (2008). *The internal structure of noun phrases in Chinese*. Crane Publishing Company.

Huang, C.-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. thesis, dissertation, MIT.

Huang, C.-T. J., Y.-h. A. Li & Y. Li (2009). *The syntax of Chinese*. Cambridge University Press Cambridge.

Jiang, J. L. (2009). Monotonicity and measure phrases in chinese. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, Chic Ling Society, vol. 45, pp. 303–317.

Jiang, L. (2012). *Nominal arguments and language variation*. Ph.D. thesis, dissertation, Harvard.

Kennedy, C. (2012). The composition of incremental change. Demonte, V. & L. McNally (eds.), *Telicity, Change, State: A Cross-categorical View of Event Structure*, Oxford University Press, Oxford.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. *Semantics and Contextual Expression* 75, p. 115.

Krifka, M. (1998). The origins of telicity. *Events and Grammar* 197, p. 235.

Li, X. (2013). *Numeral Classifiers in Chinese: The Syntax-Semantics Interface*. Walter de Gruyter.

Liao, W.-W. R. (2014). Process-related durative phrases as numeral-classifier phrases in chinese. *Taiwan Journal of Linguistics* 12:2, pp. 59–80.

Piñón, C. (2008). Aspectual composition with degrees. *Adjectives and adverbs: Syntax, semantics, and discourse* pp. 183–219.

Sybesma, R. (1999). *The Mandarin VP*. Kluwer Academic Publishers, Dordrecht, [Netherlands].

Tang, C.-C. J. (1990). *Chinese phrase structure and the extended X'-theory*. Ph.D. thesis, Cornell University.

Tenny, C. (1992). *The Aspectual Interface Hypothesis*.

Labeling and two types of null operators in English

Ryosuke Hattori

This paper investigates null operator (Op) movement in English clausal comparatives, temporal clauses and *tough* constructions. It claims that the difference among them, which concerns the clause-boundedness of Op movement in *tough* constructions, can be explained by assuming that a Complex Null operator (CNO) is involved in such constructions based on Hicks's (2009) smuggling approach while the null operators in other environments (comparatives, temporal clauses) are the simple head Ops. The Activation Condition (Chomsky 2000) and the Labeling Algorithm (Chomsky 2013) are appealed to in implementing feature-checking between the CNO and CP, which has a freezing effect on the operator in the embedded clause (Rizzi 2006, 2007). It is also argued that the simple Op keeps moving to project itself in the complement positions of certain prepositions which can assign Inherent Case through the head-complement relation.

1. Introduction

The *tough* constructions in English are represented by the following sentences in (1), where a missing object in the embedded infinitival clause is interpreted as coreferential with the subject in the main clause.

(1)   a. John is easy to please.
       b. This book is tough to read.

Since Chomsky (1977), it is assumed that the null operator (Op) is moved to the Spec-CP (though there are alternative views, e.g. Rosenbaum 1967, Postal 1971, Postal and Ross 1971, Rosenbaum 1967, Brody 1993, Messick 2012), as shown in (2).

(2)   a. John$_i$ is tough [$_{CP}$ Op$_i$ [$_{TP}$ PRO to please t$_i$]].
       b. This book$_i$ is tough [$_{CP}$ Op$_i$ [$_{TP}$ PRO to read t$_i$]].

Stowell (1986) shows (also noted by Chomsky 1973, Stowell 1987, Browning 1987, Cinque 1990; among others) that Op movement in *tough* constructions is clause-bounded (movement from the subject/object position of a finite clause is impossible), as in (3).

(3)   a. *Betsy$_i$ is easy [Op$_i$ [PRO to expect [t$_i$ fixed the car]]].
       b. *John is easy [Op$_i$ [PRO to believe [t$_i$ kissed Mary]]].
       c. ??This car is hard [Op$_i$ [PRO to claim [[Betsy fixed t$_i$]]]].
       d. ??That language is impossible [Op$_i$ [PRO to say [[Greg will learn t$_i$]]]].
                                                           (Stowell 1986: 477)

However, in other structures where Op movement is involved, e.g. comparative/temporal clauses, the extraction of the Op from an embedded finite clause is possible (Bresnan 1975, Larson 1990) as in (4).

(4)   a. Mary read more books [$_{PP}$ than [$_{CP}$ Op$_i$ everyone thinks [$_{CP}$ t$_i$ Tom believes [$_{CP}$ t$_i$ that it
          is said [t$_i$ that John read t$_i$]]]]].
       b. I saw Mary in New York [$_{PP}$ before [$_{CP}$ Op$_i$ she claimed [$_{CP}$ t$_i$ that she would arrive
          t$_i$]]].
          (Interpretation: 'prior to the time t that she alleged would be the time of her arrival.')

The clause-bounded effect in (3) is explained in the previous works based on the Empty Category Principle (e.g. Chomsky 1981); however, this effect has not been discussed in recent years. This paper proposes a way of capturing the clause-bounded nature of *tough* constructions in English within the current minimalist framework. Section 2 explores the previous accounts of the clause-bounded nature of English *tough* constructions and points out problems with them. Section 3 proposes a new analysis based on Hicks (2009) to capture the clause-bounded effect, also appealing to the Labeling theory from Chomsky (2013). Section 4 compares *tough* constructions with other Op-movement constructions and claims that there are two types of null operators which behave differently in terms of labeling. Section 5 concludes the paper.

2. Previous accounts
2.1. Stowell (1986)

Stowell himself explains the phenomenon in (3) based on the Empty Category Principle under the proper government relation (e.g. Chomsky 1981). For instance, the ungrammaticality of the (3a, b) is explained by assuming that the Op trace is not properly governed, since it is not governed by a θ-governing verb and lacks an antecedent governor assuming that a null category (=Op) may not function as an antecedent governor. On the other hand, for (3c, d), where the Op is extracted from the object positions, he claims that there is no successive cyclic movement of the Op to the Comp position because that would cause an ECP violation. Instead, the Op is forced to move directly (one-step), which incurs a pure Subjacency violation (without ECP). It is well known that violations of Subjacency yield somewhat weaker effects on acceptability which explains the slight acceptability difference between (3a,b) and (3c,d).

   However, I must note that under the current minimalist framework, the ECP analysis cannot remain as is, as the notion of government has been abandoned, which makes it necessary to reformulate Stowell's account in a different way.

2.2. Takahashi (1997)

Takahashi (1997) explains the ungrammaticality of (3a, b) by appealing to the notion of move F. The idea (from Chomsky 1995) is that a movement should take the form of feature movement (or move-F) unless violating (5b). Under this analysis, every instance of overt category movement including wh- or NP-movement involves pied-piping of extra features for PF interpretation, while covert LF movement should be in the form of pure feature movement.

(5)   a. A feature to be checked carries along just enough material for convergence
           (namely, to produce legitimate PF and LF outputs).
       b. Isolated features and other scattered parts of words are not subject to rules of the
           phonological component, violating Full Interpretation at PF (which makes derivations
           not converge, namely, crash at PF).

(Takahashi 1997: 183)

According to the Move-F hypothesis, a sentence like (3b) is represented as below, where OP stands for the null operator and $[F_1, \ldots , op, \ldots , F_n]$ for the set of features of the OP.
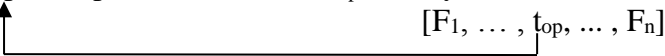
(6)   John is easy [$_{CP}$ [op]-Comp [PRO to believe [OP kissed Mary]]].
                                        $[F_1, \ldots , t_{op}, \ldots , F_n]$

Here the null operator is in Spec-IP and only the op feature moves out to the Spec-CP position. We are then dealing here with extraction out of a subject in Spec-IP, which is independently known to be impossible (the subject condition):

(7)   a. *Who did you believe [[a picture of t] was on sale]?
       b. $^?$*Who did you believe [[a picture of t] to be on sale]?

Movement of the Op feature from this position is prohibited. Thus, the ungrammaticality of the sentences in (3a, b) is explained by the Move-F hypothesis.
    This analysis, however, does not explain the ungrammaticality of (3c, d), since they would be predicted to be grammatical, with the op feature moving from the object position.

(8)   a. ??This car is hard [op-Comp [PRO to claim [t_{op} [Betsy fixed OP]]]].
                                        $[F_1, \ldots , t_{op}, \ldots , F_n]$

       b. ??That language is impossible [op-Comp [PRO to say [t_{op} [Greg will learn OP]]]].
                                        $[F_1, \ldots , t_{op}, \ldots , F_n]$

Since the movement from this position is possible as shown below, the position should not be a barrier for Move-F.

(9)   Who$_i$ did you believe Betsy saw [pictures of t_i]?

Thus, Takahashi's (1997) analysis only captures the subject extraction cases in (3a, b), not the object extraction cases in (3c, d). In the next section, I show that labeling algorithm from Chomsky (2013) together with a new analysis from Hicks (2009) provides an explanation for the clause-boundedness effect in (3a-d).

3. Explanation
3.1. Complex null operator

The analyses of *tough* construction have encountered difficulties with at least one of the core theoretical concepts of Case, locality constraints, and θ-role assignment. For instance, the raising analysis of the *tough* subject from the embedded object position by A-movement (see e.g. Rosenbaum 1967) leads to a problem with Case assignment i.e. the *tough* subject should not be able to avoid accusative case assignment in the embedded clause. Also, this approach has to explain why the A-movement here can skip another subject position (i.e. the subject of the infinitival clause, PRO). On the other hand, an account based on A'-movement of a null Op (Chomsky 1977) assumes that the *tough* subject is base-generated in situ, which apparently leaves it without a θ-role, since the *tough* predicate is assumed to not assign a θ-role to its subject. This is shown by the grammaticality of the *tough* sentences with expletive/sentential subjects in (10), which are contrasted with other configurations with missing object, as with *pretty* in (11).

(10) a. It is tough to please linguists.
      b. To please linguists is tough.

(11) a. These flowers are pretty to look at.
      b. *It is pretty to look at these flowers.
   c. *To look at these flowers is pretty.

Thus, the A'-movement analysis has to explain how a single θ-role assigned by the embedded verb is apparently 'shared' between two arguments, i.e. the null operator in the infinitival clause and the *tough* subject. Postal (1971), Postal and Ross (1971), Rosenbaum (1967) and Brody (1993), on the other hand, propose a composite A/A'-movement analysis by claiming that A'-movement of the *tough* subject to the embedded Spec, CP is followed by A-movement to the matrix subject position as shown below.

(12)   He$_i$ is easy [$_{CP}$ t$_i$ [$_{TP}$ PRO to please t$_i$]].

However, the problems of this approach are the Case mismatch of the subject (Accusative as an object of the embedded infinitival verb vs. Nominative as a matrix subject) and that movement into an A'-position that is followed by A-movement of the same object is typically assumed to be banned as an Improper Movement configuration (See Bruening 2001 and Svenonius 2004).

   Hicks (2009) proposes a new analysis which incorporates both A-movement and A'-movement without the above problems of the previous approaches, using smuggling (Collins 2005a, b). He claims that a null operator in *tough* constructions is a *wh*-phrase with more
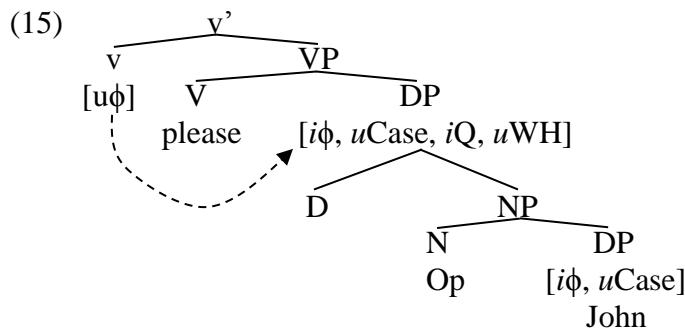
complex internal structure than is typically assumed, i.e. a complex DP with the internal DP as the *tough* subject as shown below.

(13)

DP1$^{[i\phi, u\text{Case}, iQ, u\text{WH}]}$
- D
- NP
  - N
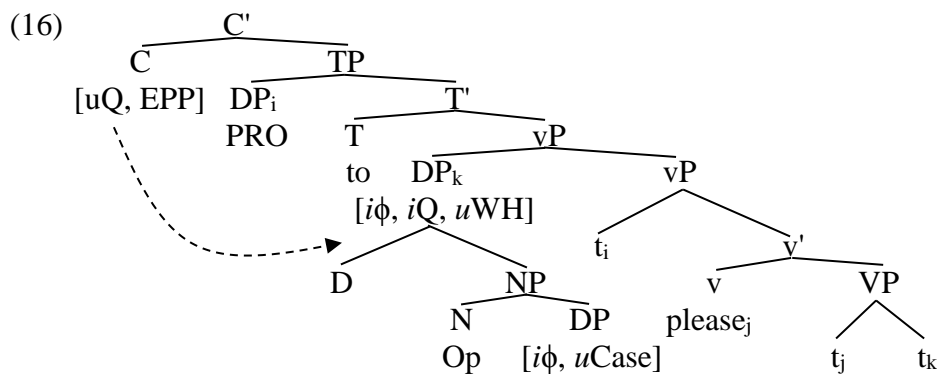   - Op
  - DP2$^{[i\phi, u\text{Case}]}$
    - John
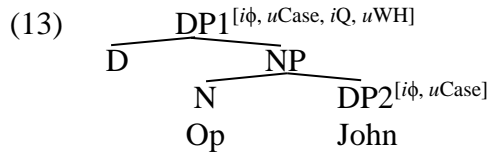
Based on this complex null operator (henceforth, CNO) analysis, the derivation of the *tough* sentence in (14), for example, proceeds as shown in (13)--(15).

(14)   John is tough for us to please.

First, the CNO merges with the V *please* as an object and the patient θ-role from the verb is assigned to the whole complex DP. Second, the derived VP is merged with v, and the complex null operator enters into ϕ-feature agreement with v, [uϕ] on v being the relevant probe. As a reflex of ϕ-feature agreement, v checks [uCase] on the CNO at this point.

(15)

v'
- v [uϕ]
- VP
  - V please
  - DP [*i*ϕ, *u*Case, *i*Q, *u*WH]
    - D
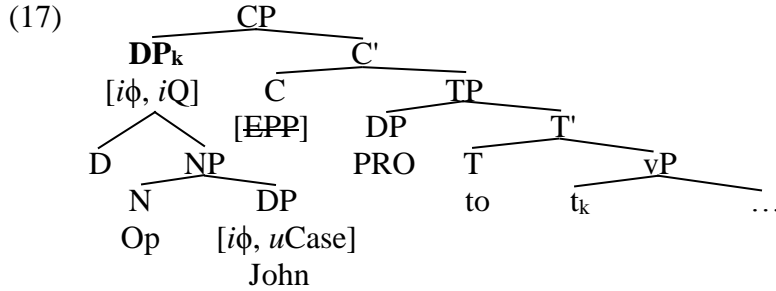   - NP
      - N Op
      - DP [*i*ϕ, *u*Case] John

As shown in (16), after V-to-v movement of *please* and the merger of PRO as the external argument in Spec, vP, the CNO must move to the phase edge (outer vP-Spec) since it bears [iQ, uWH] feature[1], where crucially the operator pied-pipes the inner DP *John*, allowing [uCase] on it to escape. The null operator therefore serves to 'smuggle' the *tough* subject. The PRO, then, moves into Spec, TP of the embedded clause, and the C is merged with [uQ] which is checked with [iQ] on the CNO while the [uWH] is checked as a reflex.

(16)

C'
- C [uQ, EPP]
- TP
  - DP$_i$ PRO
  - T'
    - T to
    - vP
      - DP$_k$ [*i*ϕ, *i*Q, *u*WH]
        - D
        - NP
          - N Op
          - DP [*i*ϕ, *u*Case]
      - vP
        - t$_i$
        - v'
          - v please$_j$
          - VP
            - t$_j$
            - t$_k$

---

[1] Cf. Bošković 2007, where the presence of an uninterpretable feature induces movement to a phasal edge.

John

As shown in (17), the [EPP] on C then drives movement of the CNO into the phase-edge position, allowing the unchecked [uCase] on *John* to escape. At this point the remaining interpretable features in the CNO are inactive.

(17)

```
                        CP
        DPₖ                        C'
      [iφ, iQ]      C              TP
                   [EPP]      DP          T'
      D     NP     PRO     T          vP
         N      DP         to     tₖ          …
        Op    [iφ, uCase]
              John
```

Finally, when the main clause T merges into the structure, T, which has [uφ], probes for [iφ]. As a reflex of φ-agreement, a nominative case value is assigned to the goal *John*, which moves to Spec, TP to satisfy [EPP], and its [uCase] is checked.

3.2. Labeling and freezing

In this section, I apply Chomsky's (2013) system to this analysis. Chomsky (2013) argues for the following Labeling algorithm: when a head (i.e. a non-branching element) and a phrase (i.e. a branching element) merge, the head projects as in (18a); when two phrases are merged, a shared feature of the two phrases is projected as in (18b) or if one of the phrases is a trace it gets ignored and the other phrase projects as in (18c).

(18) a.   X   b.   f   c.   YP
      X   YP   XP[f]   YP[f]   t   YP
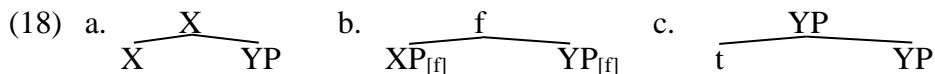
Successive cyclic movement of a *wh*-phrase is then forced since in an intermediate position where a whP and a CP are merged, creating {α whP, CP}, the syntactic object α cannot be labeled (due to the lack of feature-sharing) unless whP raises up further, its trace being ignored so that α is labeled as CP based on (18c). In some cases, a whP shares a feature with a CP and thus stops raising. For instance, this happens in an indirect question where a wh-phrase moves to merge with the CP in the complement position of *wonder*.

(19)   they wondered [α in which Texas city [β C [JFK was assassinated]]].

(Chomsky 2013: 45)

Here, the α is of the form {α whP, CP}, but whP does not raise. Based on the algorithm in (18b), the most prominent feature of whP and of CP, namely the interrogative feature Q (Cable 2007, 2010; Narita 2011), being shared, it projects as the label of α. Thus, the Freezing effect (Rizzi 2006, 2007) is now interpreted based on the labeling theory, i.e. an element undergoing A'-movement gets frozen for further movement when it participates in feature-sharing for labeling.

I claim that Op involved in *tough* constructions is an XP in terms of labeling in (18b), based on the complex null operator (CNO) analysis of Hicks (2009). Recall that he assumes that when the CNO merges with the V as an object, the patient θ-role is assigned to the whole complex DP1, and after the CNO merges with a CP, the inner DP2 is smuggled (Collins 2005a, b) into the matrix subject position[2] without being assigned accusative Case. What is important here, which Hicks does not discuss, is that assuming this CNO analysis, the CNO (=DP1) shares the Q feature with a CP when it is internally merged with it in (20).

(20)
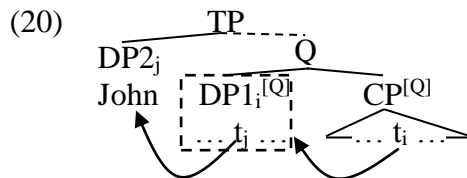
Thus, the CNO gets frozen after it A'-moves and merges with a finite clause, with the shared feature projecting as the label, just as in the case of indirect question. The ungrammaticality of (3), therefore, can be explained based on the labeling theory in Chomsky (2013).

4. Consequences
4.1. Op projecting

Since there is no freezing effect in Op-movement in comparatives or temporal clauses as in (4) (repeated below), the above analysis of Op involved in *tough* constructions has the consequence that there is no feature-sharing in Op movement in comparative or temporal clauses.

(4)   a. Mary read more books [$_{PP}$ than [$_{CP}$ Op$_i$ everyone thinks [$_{CP}$ t$_i$ Tom believes [$_{CP}$ t$_i$ that it is said [t$_i$ that John read t$_i$]]]]].
       b. I saw Mary in New York [$_{PP}$ before [$_{CP}$ Op$_i$ she claimed [$_{CP}$ t$_i$ that she would arrive t$_i$]]].
          (Interpretation: "prior to the time t that she alleged would be the time of her arrival.")

I argue that the Op here is a head (=X) in terms of Labeling and thus projects as a label after movement as in (18a) in comparative or temporal clauses, in contrast with the CNO in *tough* constructions.

Two possibilities have been argued to exist when an operator like element (OP) undergoes movement. OP can either not project as in (21), where OP is the Spec-CP, or merge with CP and project itself as a label as in (22).

(21)

(22)

$$OP \begin{cases} \text{Overt e.g. } wh \\ \text{Null Op} \end{cases}$$

---
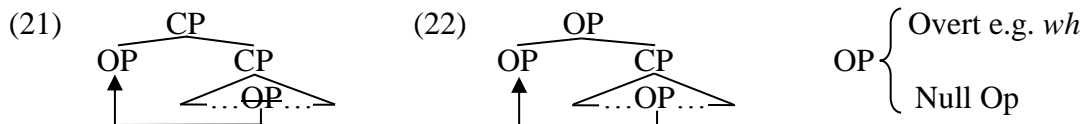
[2] Here I assume the DP2 is subextracted through the edge of the DP1 phase.

The case from (21) is found in wh-questions as shown below, where it is standardly assumed that a *wh*-phrase moves to the Spec-CP, but does not project.

(23)   [$_{CP}$ What$_i$ [$_{C'}$ did John say t$_i$]]?

The second case in (22) is instantiated by English free relatives (Donati 2006, Cecchetto and Donati 2010, 2011, 2015); the overt operator (= wh) in (24) moves to merge with CP; it is assumed to project since the DP interpretation is obtained (see the works cited above).

(24)   I read [$_{DP}$ [$_D$ what]$_i$ [$_{CP}$ you read t$_i$]].   (Cecchetto and Donati 2015; 1)

Here, the wh-word moves as the D head[3] in this context and provides a label when it merges with a CP, so that resulting structure *what you read* becomes a DP. Looking at null operators in this regard, the Op in comparative/temporal clause seems to project as a label in the same way, since crucially we get the same DP interpretation of the relevant clause in such cases (e.g. 'the amount in which John read books.' or 'time of Mary's arrival').

4.2. Comparative Clauses

A number of authors have claimed that comparative clauses should be analyzed as a kind of free relatives (Bracco 1980, Donati 1997, Larson 1987, Grosu 1994, etc.). The idea that comparative clauses are similar to free relatives can be attributed to Larson (1987), who claims that adjectival free relatives such as *however tall Bill is* are not in fact relatives but 'free comparatives'. This idea comes from intuitions about the meanings of examples like (25).

(25)   a. John will grow [however tall his father did].
          b. I'll word my letter [however carefully you word yours].

These sentences can be paraphrased exactly by means of explicit comparatives as in (25').

(25')   a. John will grow as tall as his father did.
          b. I'll word my letter as carefully as you word yours.

Donati (1997) claims that movement of the bare head Q in comparatives is desirable, since comparative clauses obey a strong anti pied-piping condition, as shown below, where the pied-piping of an extra element (i.e. *candies* or *to*) is prohibited.

(26)   a. *Mary ate more cookies than [e candies]$_i$ she ate [e]$_i$.
          b. *I talked to more people than [to e]$_i$ I wrote [e]$_i$.

She claims that the nominal/prepositional associate of the moving quantifier here is not pied-piped, suggesting involvement of head movement. It is possible to interpret these data as null operator being incompatible with pied-piping. However, even in Italian, where the moved Q is overt (= *quanto*), the pied-piping is impossible as in (27).

---

[3] Here the element is actually non-branching and thus ambiguous between head and phrase.

(27) a. *Ho   mangiato più biscotti di [quante caramelle]$_i$
          Have-1SG eaten   more cookies than WH-PL candies
          abbia   mangiato [e]$_i$.
          have-1SG eaten
          'Intended: I ate more cookies than candies I ate.'
      b. *Ho   parlato a più   persone di [a quante]$_i$ abbia
          Have-1SG talked to more persons than to WH-PL have-1SG
          scritto [e]$_i$.
          written
          'Intended: I talked to more people than to I wrote.'   (Donati 1997: 153-154)

This shows that the movement involved in comparative clauses is different from normal wh-movement; it must be a bare element which moves, just like in free relatives. In fact, pied piping is also bad in free relatives; both in English and Italian, where *wh*-element pied-pipes a preposition governing it, as shown below.

(28) a. Mary considers [what]$_i$ Mary is talking about [e]$_i$.
      b. *Mary considers [about what]$_i$ Mary is talking [e]$_i$.

(29) *Mario ascolta [di   quanto]$_i$ sta parlando Maria [e]$_i$.
       Mario listens about what   is talking   Mary
       'Mario listens to what Mary is talking about.'

Further support for the claim comes from a difference between free relatives and headed relative clauses. Citko (2008) shows that free relatives in English as in (30a) disallow the complementizer *that* while headed relatives allow it as in (30b).

(30) a. We'll hire whom (*that) you recommended to us.
      b. We'll hire the man (that) you recommended to us.

Crucially, comparative clauses pattern with free relatives in this respect as shown below, where *that* is not allowed in the complement of *than*, suggesting that what is involved here is a free relative like structure rather than a headed relative.

(31) John is taller than (*that) Mary is tall.

Furthermore, Donati (2006) points out that clausal comparatives in general should involve *wh*-head movement and not the movement of a wh-phrase, referring to Romanian data from Grosu (1994). In this language, a wh-element *cit* 'how' appears as the overt operator in comparatives.

(32) a. Ion e şi   mai puternic decît[4] a   fost tatăl   lui.
          Ion is and more strong   of-how has been father-L his
          'Ion is even stronger than his father was.'

---

[4] Grosu argues that *cit* combines with *de-* in the lexicon (Grosu 1994; p. 208).

     b. Ion e tot atît   de puternic pe cît   a fost tatăl   lui.
        Ion is all so-much of strong   as how has been father-L his
        'Ion is just as strong as his father was.'

Moreover, the same *wh*-element allows pied-piping in interrogatives or exclamatives.

(33) a. Cît   de frumoasă e Maria?
       how of beautiful is Maria
       'How beautiful is Maria?'
     b. Cît   de frumoasă e Maria!
       how of beautiful is Maria
       'How beautiful Maria is!'

      c. Cît   e Maria de frumoasă?
       how is Maria of beautiful
       'How beautiful is Maria?'
     d. Cît   e Maria de frumoasă!
   how is Maria of beautiful
       'How beautiful Maria is!'

However, the very same *wh*-element *cit* 'how' does not allow pied-piping in comparatives.

(34) a. *Maria e cu   mult mai desteapta [decît   de frumoasa] e Zamfira.
       Maria is with much more clever   of-how of beautiful is Zamfira
       'Mary is much more clever than beautiful Zamfira is.'
     b. Maria e cu   mult mai desteapta [decît] e Zamfira de frumoasa.
       Maria is with much more clever   of-how is Zamfira of beautiful
       'Mary is much more clever than Zamfira is beautiful.'   (Grosu 1994)

This suggests that the degree operator in Romanian moves as a head in comparatives, not as a phrase.
   Considering the above discussion, i.e. the free relative like interpretation/behavior, the comparative clause in English is analyzed as in the following.

(35) a. This table is taller than [DegP [Deg Op]i [CP that door is wide [DegP [Deg t]i]].
     b.
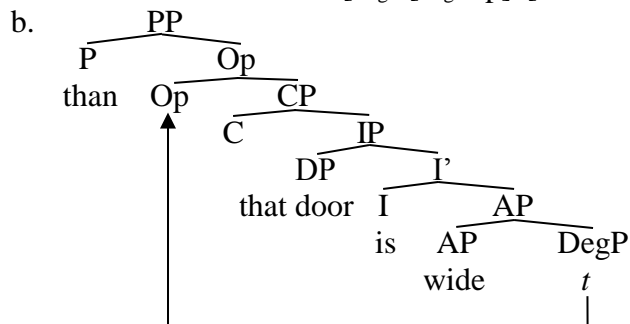
Here the DegP is a null equivalent of *degree/amount*[5], which moves as a Deg head (actually being non-branching, it is an ambiguous head/phrase) from the sentence-final position and internally merges with the CP, and then projects itself as a label.[6] What is important is that the Op involved here is non-branching, hence it is both a phrase and a head in the bare phrase structure theory (Chomsky 1995).

Based on this Op projection analysis, the bare clause of comparatives like *that door is wide* in *This table is taller than that door is wide* is analyzed as a projected Op (i.e. [$_{Op}$ Op$_i$ that door is wide t_i]). Although the projected Op has a DP like interpretation, i.e. 'the degree to which that door is wide', the case in question has a very restricted distribution, unlike free relatives (with overt operator projecting): the Op projection seen in comparatives is possible as the complement of a P *than*, while it cannot appear in the subject position as in (36a) or the complement of a V as in (37a). On the other hand, an NP overtly headed by *degree* can appear in such positions as in (36-37b).

(36) a. *[$_{Op}$ Op$_i$ that door is wide t_i] is surprising.
     b. [$_{NP}$ The degree (to which that door is wide)] is surprising.

(37) a. The length of this table exceeds *[$_{Op}$ Op$_i$ Mary is tall t_i].
     b. The length of this table exceeds [$_{NP}$ the degree (to which Mary is tall)].

In light of this, I propose a generalization regarding the distribution of the projected Op, i.e. it cannot appear as a subject or a complement of a verb, as summarized in (38).

(38) The Op Generalization
     a. null Op projection can be the complement of certain Ps like *than* in
        comparatives.
     b. null Op projection cannot be the complement of a verb.
     c. null Op projection cannot be a subject.

(38') a. ✓

---
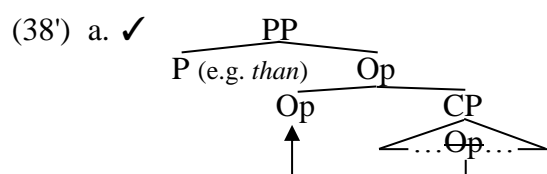
[5] Following Izvorski (1995, 2000) who proposes that a degree/amount phrase (null counterparts of the expressions like *in what quantity/to what extent*) moves out of a sentence-final adjunct position as shown below.
(i) a. …than [$_\varphi$ in what quantity]$_i$ we have books t_i.
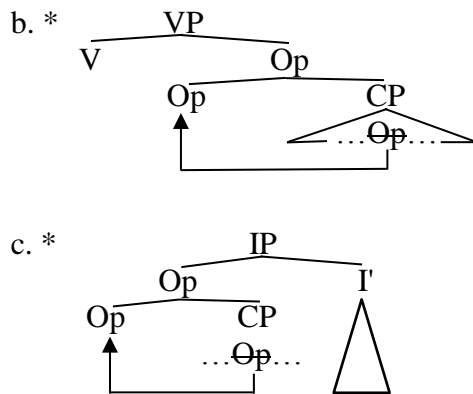   b. …than [$_\varphi$ to what extent/degree]$_i$ it is wide t_i.
[6] Donati (1997, 2006) assumes a slightly different structure for Italian comparatives as shown below, where the overt operator moves as a determiner-like Q head to the C head position and Q and C both project when internally merged with the IP, turning the clause into a complex nominal.
(i) [$_{PP}$ [$_P$ di] [$_{CP/QP}$ [$_{C/Q}$ quanti$_i$] [$_{IP}$ … t_i …]]].
            than   x-much
Izvorski (2000), on the other hand, does not assume that the moved element projects, as in (ii). She claims that the NP denotation is obtained through semantic type-shifting from (iii-a) to (iii-b).
(ii) [$_{PP}$ [$_P$ than] [$_{CP}$ wh$_i$ … [$_{XP}$ [$_{X'}$ X t_i]]]].
(iii) a. …than [$_{FR}$ [to what degree]$_i$ Ann was sad t_i before].
      b. …than [$_{NP}$ the degree, to which Ann was sad before].

b. *

```
            VP
      V           Op
          Op           CP
           ↑      ⟨ …Op… ⟩
           └──────────┘
```

c. *

```
              IP
        Op           I'
    Op      CP       △
     ↑    …Op…
     └──────┘
```

4.3. Temporal Clauses

Turning to temporal clauses, Larson (1990) observes the following paradigm regarding temporal prepositions. First, clausal PPs[7] headed by temporal prepositions like *before, after, since* and *until* show ambiguity in their interpretation (Geis 1970).

(39) a. I saw Mary in New York [PP before [CP1 she claimed [CP2 that she would arrive]]].
       b. I encountered Alice [PP after [CP1 she swore [CP2 that she had left]]].
       c. I can't leave [PP until [CP1 John said [CP2 I could leave]]].
       d. I haven't been there [PP since [CP1 I told you [CP2 I was there]]].

The temporal clause in (39a) can be interpreted as 'before she made a certain claim' or 'prior to the time t that she alleged would be the time of her arrival.' Larson claims that the lower interpretation (the latter interpretation referring to the time of 'arrival') arises via null Op movement as in (40b), since the ambiguity disappears with a subjacency violation (Complex NP Constraint), as in (40a). I assume (following Larson) that Ps can pick up the temporal reference of the highest CP in its complement. Then, the local construal (the first reading noted above, referring to the time of 'claim') is obtained without Op-movement as in (40a).

(40) a. I saw Mary in New York [PP before [she made [NP the claim [that she had arrived]]].
           ✓ "Before she made a certain claim"
           * "prior to the time t that she alleged would be the time of her arrival."
       b. I saw Mary in New York [PP before [CP1 Op$_i$ she claimed [CP2 t$_i$ that she would arrive t$_i$]]].

On the other hand, other English clausal Ps, e.g. *although, because, in case, unless*, and *while*, are not ambiguous in this manner.

(41) a. I still respect John [PP although [he claims [that he killed his mother]]].
       b. I visited New York [PP because [Mary dreamed [that Max was there]]].

[7] I refer to a phrase in which a P appears to take a clausal complement as 'clausal PP'.

c. I won't visit New York [PP unless [Bill promises [Mary will be there]]].
d. I won't visit New York [PP in case [Bill says [Mary is there]]].
e. I didn't see Mary in New York [PP while [CP1 she said [CP2 she was there]]].

For example, the expected two readings in (41a) would be paraphrased as 'despite John's claiming that he killed his mother' and 'despite what John claims, viz., that he killed his mother', but only the former reading is possible where the concession can only be understood as being made for John's claim. Therefore, I assume (following Larson 1990) that there is no Op movement available in the complements of these prepositions (*although, because, in case, unless, while*). Larson claims that the crucial difference between the two groups is that the prepositions of the former group can take both DP and clausal complements.

(42) a. {before / after / since / until} {John arrived / that day}

b. while {John slept / *that day}

c. {although / because / unless / in case} {Mary walked out on Max / *that fact/reason/eventuality}

The following explanation can be given for Larson's paradigm, based on the assumption that Op can project when it is internally merged with a CP. In principle, after Op moves, either Op or CP can project. However, the former option is possible only with the Ps that can take DP complements, hence only the Ps from (42a). It is not an option for the Ps from (42b-c). The structure from (39') is then possible only for Ps from (42a).

(39') a. I saw Mary in New York [PP before [Op Op$_i$ she claimed [t$_i$ that she would arrive t$_i$]]].
b.
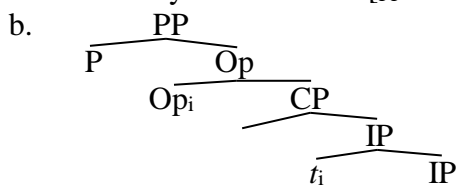
(43) a. {before / after / since / until} {Op / DP / CP}

b. {while / although / because / unless / in case} {*Op / *DP / CP}

In other words, the set of Ps from (42a) can take an Op projection as their complement, while the Ps from (42b-c) cannot. Therefore, the movement from a deeply embedded clause is

possible only when the prepositions can take the Op projection as their complements as in (39'b). This leads to the conclusion that long-distance movement of the null operator, hence the long-distance construal, is possible only if Op projects after movement, hence with the Ps from (43a), but not if there is no Op movement, hence not with the Ps from (43b-c).

Further, the relevant Op projection here, i.e. '[$_{Op}$ Op$_i$ she claimed [t$_i$ that she would arrive t$_i$]]]' in (39'a), cannot appear in an argument position (subject or object) of a verb without an overt head as in (44c, d) while the NP *the time* can appear in the same environments as in (44a, b).

(39') a. I saw Mary in New York [$_{PP}$ before [$_{Op}$ Op$_i$ she claimed [t$_i$ that she would arrive t$_i$]]].

(44)   a. [The time] was too late.
       b. John misunderstood [the time].
       c. [*(The time when) she claimed that she would come] was too late.
       d. John misunderstood [*(the time when) she claimed that she would come].

This confirms that the Op projection cannot be in the subject position or selected by a verb as its complement (the Op Generalization in (38)). The remaining questions that arise here are why the Op keeps moving until it projects in comparative/temporal clauses, and why it has to be projected in the complement position of certain Ps e.g. *than* or *before*.

4.4. Deducing the Op Generalization

This subsection proposes an account of the Op generalization in (38). Chomsky (2000) claims that when there is a probe seeking a goal, triggering its movement, a goal must bear some uninterpretable feature to be visible for the probe. This is referred to as the Activation Condition. Schematically, the feature-checking system functions as shown below (i = interpretable feature, u = uninterpretable feature).

(45)   X (probe) Y(goal)
       uF   iF
      EPP   uK

Here, the goal Y is targeted by the probe X and the feature F is deleted under match (Agree). As a reflex of F feature checking relation, the uninterpretable feature K of Y is checked off. For example, in (46), the C and *what* are involved in a wh-feature checking relation, and the uninterpretable Q feature of the Goal *what* made it visible for the probe, and thus *what* moves to Spec-CP. The Q feature is, in turn, checked off after the movement.

(46)   I wonder what$_i$ C   Mary bought t$_i$.
             iwh ~~uwh~~
             ~~uQ~~ ~~EPP~~   (Bošković 2007: 599)

When it comes to agreement and Case, Case makes DP/NP active for agreement/movement, and is checked off as a reflex of agreement/movement, as shown below. φ-feature licensing is then necessary for Case to be checked by a verb, since structural Case licensing is a reflex of φ-feature checking (for Chomsky).

(47) v(probe) NP(goal)
      ~~uφ~~   iφ
      ~~uK~~

I claim that the null Op lacks φ-features but has an uninterpretable Case feature to be checked off. Assume that in English the Activation Condition holds. After Op moves, Op projects and the projected Op needs to be Case-checked. Since Op has no φ-features, there will be no φ-feature checking relation between V and Op and thus V cannot assign structural Case to it.

(48) V(probe) Op(goal)   cf. (47)
      uφ   [dashed box]
      uK

(49) P(probe) Op(goal)
      ~~uK~~

On the other hand, prepositions are often assumed to assign inherent Case. Furthermore, inherent Case is standardly assumed to be dissociated from agreement. I claim, then, that prepositions like *than* or *before* can assign inherent Case, which needs no φ-features for checking, and thus the Case of the projected Op can be checked without establishing a feature checking relation. This is precisely the reason why the Op projection can appear in the complement of some Ps (which can assign inherent case) and not of Vs. The Op has to keep moving to project itself as a label in order to be in the prepositional complement position where it gets its uK checked by being assigned inherent Case by the comparative preposition *than* or temporal prepositions like *before*.

(50)
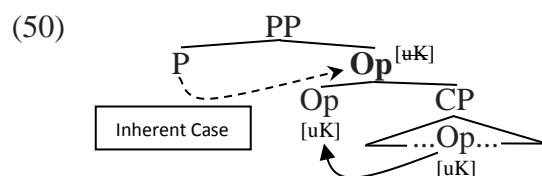
This explains the non-clause-bounded nature of these constructions, i.e. comparatives and temporal clauses in English; Op never projects when it is merged with a CP in the most embedded clause where it cannot be Case licensed, thus it keeps raising. This account of Case-licensing is in the spirit of Bošković (2007), where it is claimed that the need for a NP/DP to check Case can drive the movement of the NP/DP.

(51) a. Mary read more books **than** [Op$_i$ everyone thinks Tom believes [t$_i$ that it is said t$_i$
         that John read t$_i$]].   [~~uK~~]

      b. I saw Mary in New York [$_{PP}$ **before** [Op$_i$ she claimed [t$_i$ that she would arrive t$_i$]]].
                                                                     [~~uK~~]

Now, inherent Case is often assumed to be assigned under both θ-role assignment and in a head-complement relation. The above discussion indicates that the second requirement (a head-complement relation) is the only pre-requisite for inherent Case assignment (since the preposition does not appear to assign a θ-role to the element it Case-marks in the relevant cases). This is actually in line with Franks (1994), who argues that a head-complement relation but not necessarily θ-assignment is needed for inherent Case. In particular, Franks provides a number of arguments that the genitive case assigned by numerals to its NP complement, as in (52), is an inherent case in Serbo-Croatian.

(52)   On kupuje [pet [kola]].
       he buys   five   cars$_{GEN}$
       'He buys five cars'

Here the numeral *pet* 'five', although it assigns an inherent case to its complement, obviously does not assign a θ-role to it.[8] The impossibility of Op projection in subject position also follows given the lack of the relevant head-complement relation for Case assignment here. Regarding the fact that local temporal construal is possible with all prepositions in (43), regardless of whether they allow DP complements, following Larson (1990), I assume that the local construal comes from the temporal preposition itself, i.e. no Op is present here and hence there is no issue of Op-Case licensing.

5. Conclusion

This paper investigated the null operator (Op) movement in English comparative/temporal clauses and *tough* constructions. It was claimed that the difference among them, where only *tough* constructions display clause-boundedness of Op movement, can be explained by assuming that a Complex Null operator (CNO) is involved in such constructions based on the smuggling approach by Hicks (2009) under Labeling theory by Chomsky (2013). I claim that the CNO, which is a phrase in terms of Labeling algorithm, gets frozen when it merges with a CP as the shared feature projects as the label. The paper further argued that the Op in comparative/temporal clauses is a head (i.e. non-branching), hence it projects as a label after movement. I have established a generalization regarding the distribution of the Op-projecting option. In particular, I argue that the projected Op can be the complement of certain prepositions but not the complement of a verb or a subject. To explain this generalization and the unbounded nature of Op movement in comparative/temporal clauses, I claim that the head Op in these constructions keeps moving to project itself in the complement positions of certain prepositions, e.g. *than* or *before*, which can assign Inherent Case to check off the Op's uninterpretable Case feature without establishing a φ-feature checking relation, through the Head-Complement relation, the underlying assumption being that this Op has no φ-feature.

Acknowledgements

Ryosuke Hattori
University of Connecticut
ryosuke.hattori@uconn.edu

References

Bošković, Ž. (2007). On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38, pp. 589-644.

Bracco, C. (1980). On the Island Character of Italian 'Quanto' Comparatives. *Journal of Italian Linguistics* 5, pp.19-46.

Bresnan, J. (1975). Comparative Deletion and Constraints on Transformations. *Linguistic Analysis* 1, pp. 25-74.

Brody, M. (1993). θ-Theory and arguments. *Linguistic Inquiry* 24, pp. 1-23.

Browning, M. (1987). Null Operator Constructions. Ph.D. Dissertation, MIT.

Bruening, B. (2001). Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy. Doctoral dissertation, MIT, Cambridge, MA.

Cable, S., (2007). The Grammar of Q. Ph.D. Dissertation, MIT.

Cable, S., (2010). *The Grammar of Q: Q-Particles, wh-movement and pied-piping*. Oxford University press, Oxford.

Cecchetto, C. & C. Donati (2010). On Labeling. Principle C and Head Movement. *Syntax 13*, pp. 241-278.

Cecchetto, C. and C. Donati. (2015). *On (re)labelling*. Linguistic Inquiry Monographs. Cambridge, Mass : The MIT Press.

Chomsky, N. (1973). Conditions on transformations. In *A festschrift for Morris Halle*, ed. By Stephen Anderson and Paul Kiparsky, New York: Holt, Rinehart and Winston, pp. 232-286.

Chomsky, N. (1977). On Wh-Movement. In *Formal Syntax.* Peter Culicover, Thomas Wasow, and Adrian Akmajian (eds.), New York: Academc Press, pp. 71-132.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht:Foris.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalistin quiries: The framework. In *Step by step: Essays on minimalismin honoro of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, Cambridge, Mass.: MIT Press, pp. 89-155.

Chomsky, N. (2013). Problems of projection. *Lingua* 130, pp. 33-49.

Cinque, G. (1990). *Types of A' -dependencies*. Cambridge, MA: MIT Press

Citko, B. (2008). Missing labels. *Lingua* 118, pp. 907-944.

Collins, C. (2005a). A smuggling approach to raising in English. *Linguistic Inquiry* 36, pp. 289-298.

Collins, C. (2005b). A smuggling approach to the passive in English. *Syntax* 8, pp. 81-120.

Donati, C. (1997). Comparative Clauses as Free Relatives: a Raising Analisis. *Probus* 9, pp. 145-166.

Donati, C. (2006). On wh-head movement. In Cheng, L.L.-S., Corver, N. (Eds.), *Wh-Movement: Moving on*. MIT, Cambridge, MA, pp. 21-46.

Donati, C. & C. Cecchetto (2011). Relabeling Heads: A Unified Account for Relativization Structures. *Linguistic Inquiry*, 42, pp. 519-560.

Franks, S. (1994). Parametric properties of numeral phrases in Slavic. *Natural Language and Linguistic Theory* 12, pp. 570-649.

Geis, M. (1970). Adverbial subordinate clauses in English. Ph.D. dissertation, MIT.

Grosu, A. (1994). *Three Studies in Locality and Case*, Routledge, London.

Hicks, G. (2009). Tough-Constructions and their Derivation. *Linguistic Inquiry* 40, pp. 535-566.

Izvorski, R. (1995). A Solution to the Subcomparative Paradox. *Proceedings of the Fourteenth West Coast*

Conference on Formal Linguistics. José Camacho, Lina Choueiri, and Maki Watanabe (eds.), Stanford: CSLI
    Publications, pp. 203-219.
Izvorski, R. (2000). Free Relatives and Related Matters. Ph.D. Dissertation, University of Pennsylvania.
Larson, R. (1987). Extraction and Multiple Selection in PP. *MIT Working Papers in Linguistics* 9, pp. 119-136.
Larson, R. (1990). Extraction and Multiple Selection in PP. *The Linguistic Review* 7, pp. 169-182.
Narita, H. (2011). Phasing in Full Interpretation. Ph.D. Dissertation, Harvard.
Postal, P. (1971). *Cross-Over Phenomena*. New York: Holt, Rinehart and Winston.
Postal, P. & J.R., Ross. (1971). Tough movement si, tough deletion no! *Linguistic Inquiry* 2, pp. 544–545.
Rizzi, L. (2006). On the Form of Chains: Criterial Positions and ECP Effects. in L. Cheng, N. Corver, eds,
    Wh Movement: Moving on. MIT Press, Cambridge, Mass., pp. 97-134.
Rizzi, L. (2007). On Some Properties of Criterial Freezing. *CISCL Working Papers - STiL – Studies in
    Linguistics*, Vol.1, pp. 145-158.
Rosenbaum, P.S. (1967). *The Grammar of English Predicate Complement Constructions*. Cambridge: MIT Press.
Stowell, T. (1986). Null antecedents and proper government. *Proceedings of NELS 16,* pp. 476-493.
Stowell, T. (1987). As so, not so as. Unpublished manuscript, University of California, Los Angeles.
Svenonius, P. (2004). On the edge. In *Peripheries: Syntactic edges and their effect*, ed. by David Adger, Cécile
De
    Cat, & George Tsoulas, Dordrecht: Kluwer, pp. 259–287.
Takahashi, D. (1997). Move F and null operator movement. *The Linguistic Review* 14, pp. 181-196.

The Atoms of Person
Limitations on Concept Formation

Jolijn Sonnaert

Even though person markers such as agreement suffixes, clitics or independent pronouns are very common crosslinguistically, many questions about the possible person referents these markers can refer to remain. Consider for example the inclusive, which refers to the referents of both first and second person, respectively speaker (called *i* in this paper) and hearer (called *u*): *iu*. In a representative sample of 370 languages, the other logically possible person referent combinations, *io* (for speaker and other) and *uo* (for hearer and other), are unattested and therefore impossible person distinctions in natural language. The literature often equates *io* and *uo* with first and second person plural, respectively. However, I show that pronominal plural is not an instance of '+ third person' or '+ *o*', but that it is in fact an instance of '+ associates' or '+ *a*', resulting in the plurals *iua*, *ia*, *ua* and *oa*. The CONCEPT FORMATION CONSTRAINT in the kite framework (see Jaspers 2012; Seuren & Jaspers 2014; Roelandt 2016) makes generalisations about lexicalisation in closed lexical fields and predicts exactly the gap in lexicalisation we see for person, namely the absence of person markers for the combinations *io* and *uo*. As such, the fact that these are impossible person distinctions is part of a broader generalisation on lexicalisation in natural language.

1. Introduction

Pronominal paradigms are typically analysed as having three person atoms, corresponding to speaker, hearer and other and being referred to by first, second and third person.[1] In this paper, I refer to these atoms as *i*, *u* and *o* respectively, following a.o. Harbour (2016) and Ackema & Neeleman (To appear). As for plural, I argue that pronominal plural is formed by adding

---

[1] Note that these atoms and their combinations are not the same as the morphosyntactic features like [± participant], [± speaker], etc. familiar from the literature (e.g. Harley & Ritter 2002; Bobaljik 2008; Harbour 2016). Instead, these atoms and their combinations are the referents that these features can select for pronominals to refer to. This paper is only concerned with the atoms and combinations, not with the morphosyntactic features.

associates (*a*) to these person atoms (as already suggested by Bobaljik (2008) and Ackema & Neeleman (To appear)) rather than by adding third person, as commonly assumed in a lot of the literature on person (e.g. Cysouw 2003; Harbour 2016). The distinction between associates and others is a crucial one, and will be motivated in Section 3.1.

The relationship between possible person atoms and morphological person markers is not one to one, since on top of first, second and third person, languages can also have a distinct inclusive person. The inclusive refers to a group consisting of both the speaker and the hearer and is commonly analysed as a combination of those. For example, Cysouw (2003) describes the inclusive as '1+2', Harbour (2016) and Ackema & Neeleman (To appear) as '*iu*'. However, languages have different ways of expressing this inclusive. Consider for example Tümpisa Shoshone below, where inclusive *tammü* refers to a *we* including the hearer, as opposed to *nümmü* referring to a *we* excluding the hearer. This latter one is also called the *exclusive* or the first person plural.

(1)   Tümpisa Shoshone (Dayley 1979)

	SG			PL		
incl				*ta-mmü*	iu(a)	(we, including you)
1	*nü*	i	(I)	*nü-mmü*	ia	(we, but not you)
2	*ü*	u	(you)	*mü-mmü*	ua	(you)
3	(demonstr)	o	(he, she, it)	(demonstr)	oa	(they)

Languages like Tümpisa Shoshone, which have one distinct pronominal to express this combination of speaker and hearer, are called *inclusive languages* (terminology from Daniel 2005), as opposed to *non-inclusive languages* such as English where the inclusive is expressed syncretically with the exclusive.

The cell for pronouns that would be both 'inclusive' and 'singular' in the paradigm in (1) is empty because the inclusive can never be singular. It necessarily refers to both the speaker and hearer. Also, in agglutinative paradigms where person and number are expressed by separate morphemes like Tümpisa Shoshone above, the inclusive takes plural rather than singular morphology: just like the other plural pronouns, it adds *-mmü*.

This is not to say the inclusive cannot have a number distinction. Some languages, i.e. *minimal-augmented languages* like Rembarrnga below, have two pronominals to refer to a combination of speaker and hearer. They distinguish a minimal from an augmented inclusive: that is, an inclusive that consists of the minimal amount of participants needed for the inclusive person (*i* and *u*), *yɨkkɨ*, and an inclusive that adds to this associates (*i*, *u* and *a*), *ngakorrbbarrah*. Note that just like the plural as opposed to the singular in Tümpisa Shoshone, the augmented persons have a number morpheme, *-barrah*, that the minimal do not.

(2)   Rembarrnga (Cysouw 2003)

	MIN			AUGM		
incl	*yɨkkɨ*	iu	(we: me & you)	*ngakorr-bbarrah*	iua	(we: me, you, & more)
1	*ngɨnɨ*	i	(I)	*yarr-bbarrah*	ia	(we, but not you)
2	*kɨ*	u	(you)	*nakorr-barrah*	ua	(you)
3	*nawɨ/ ngadɨ*	o	(he, she, it)	*barr-barrah*	oa	(they)

To summarise: we can distinguish three types of languages when considering the inclusive person: non-inclusive languages like English that express the inclusive syncretically with the exclusive, inclusive languages like Tümpisa Shoshone with an inclusive that behaves like the plural persons, and minimal-augmented languages with an inclusive that behaves like the singular other persons, i.e. minimal, and one that behaves like the plural other persons, i.e. augmented. Considering that in addition to the lexicalisations of the single person atoms (i.e. first, second and third person markers), languages can also distinguish a combination of two of these atoms (inclusive, *iu*), the question arises whether they can lexicalise even more person distinctions. The three person atoms can be combined in eight different ways (3).[2]

(3) a. (ø, expletives)   e. *iu*, inclusive
     b. *i*, first person   f. *io*, −
     c. *o*, second person   g. *uo*, −
     d. *u*, third person   h. (*iuo*, impersonals)

I use a sample of 370 languages to show that only six out of these eight possibilities can be lexicalised as simplex person morphemes: *io* and *uo* cannot be lexicalised. This is especially clear when also considering number. For a mininal-augmented language as discussed above, the six logically possible person distinctions in the two numbers would yield a 12-cell paradigm, however, for these languages only 8-cell paradigms like the one for Rembarrnga are attested, with both numbers distinct for *i, u, o* and *iu*.

The absence of exactly *io* and *uo* in person lexicalistion is predicted in the kite framework developed by Seuren & Jaspers (2014). Seuren & Jaspers look at concept formation in closed lexical fields and the logical relations that hold between the concepts. In doing so, the framework makes predictions about lexicalisation. For example, for quantifiers, it predicts that languages do not have a simplex morpheme to express the contradiction of *all*, namely *not all*, nor to express the disjunction *all or no*. This is the CONCEPT FORMATION CONSTRAINT (CFC). I will show that the same constraint applies to person.

The paper is organised as follows: I first discuss the person lexicalisations and the unlexicalisable combinations in detail in Section 3. Then I introduce the kite framework and explain how it makes predictions about lexicalisation in Section 4. In Section 5, I show how this framework predicts exactly and only the attested person distinctions and the relations that hold between them as part of a broader generalisation on lexicalisation.

2.   *The sample*

The sample collected consists of two parts. The first part is a small sample of independent pronouns for only 30 languages, collected to be representative for the genetic and geographic variation as discussed in *The World Atlas of Language Structures* (Dryer & Haspelmath 2013). This was then extended to a sample of 370 languages with all paradigms available in the typologies and analyses of a.o. Forchheimer (1953); Harley & Ritter (2002); Baerman et al. (2005);

---

[2]This paper does not discuss the absence and presence of all atoms, i.e. *ø* and *iuo*. These might correspond to expletive pronouns (ø, not referring to any person atoms) and impersonal pronouns (*iuo*, referring to all person atoms) as proposed by Ackema & Neeleman (To appear). This paper focuses only on the referential and personal use of person morphemes: the options in (3b)-(3g).

Daniel (2005); Bobaljik (2008); Cysouw (2003); Harbour (2016); Ackema & Neeleman (To appear). The sample now includes all kinds of person markers: free pronouns, clitics and agreement markers.

3.   *The lexicalisation of person*

Even though analyses of person differ on which morphosyntactic features arrange the paradigms of pronominals, all these accounts assume that person markers can refer to three different persons: a speaker i, a hearer u and an other o. Pronominals can refer to these single atoms with first, second and third person, or to a combination of atoms iu with the inclusive. In this section, I will discuss the lexicalisations of all person atoms and their combinations and show that of the latter, only the inclusive ever receives a simplex lexicalisation in language, meaning that only the inclusive can be expressed with a single person morpheme. The other two combinations (io and uo) never have a distinct person morpheme.

I start with discussing the distinction between o, third person, and a, plural, in Section 3.1 on pronominal number. Next, I discuss the person atom lexicalisations in detail in Section 3.2, to conclude this section with a note on syncretism in Section 3.3.

3.1.   *Number*

Looking at number, many analyses of person collapse third person and plural (e.g. Cysouw 2003; Moskal 2014; Harbour 2016). Consider for example the following terminology used in Cysouw (2003) to represent the person and number paradigm, where the plural persons are marked as the relevant person '+ third person'.[3]

(4)   The person and number paradigm in Cysouw (2003)

PERSON	SG	PL
inclusive		1+2(+3)
exclusive	1	1+3
addressee	2	2+3
other	3	3+3

The same goes for Harbour's (2016) person analysis, for which the ontology is shown in Fig. 1. In his system, i represents the (unique) speaker, u represents the (unique) hearer and the os represent an infinite amount of others. According to Harbour, applying the features he proposes to this ontology derives all persons attested, described in (5). The subscript o points to the possible addition of one or more 'others' for plural, the normal case o points to one 'other' for third person.

---

[3]I do not wish to imply that Cysouw is committed to the belief that third person is plural, but the notation used is fairly widespread and might invite this conclusion.
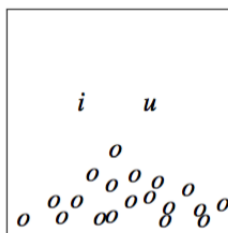
Figure 1: *The person and number ontology in Harbour (2016).*

(5) a. iu_o = inclusive
     b. i_o = first person
     c. u_o = second person
     d. o_o = third person

The fact that both plural and third person consist of this *o* is a defining characteristic of Harbour's system.

If plural and third person were the same, *uo* and *io* would refer to first and second person plural. However, the morphology and semantics of third person and plural show that it is incorrect to conflate both. In what follows, I first discuss their morphology (Section 3.1.1). I show that no language uses the same morpheme for both plural and third person, which would not be unexpected if they were indeed the same. For semantics, I follow Ackema & Neeleman (To appear) in that the reference of a plural pronoun cannot include a third person. I demonstrate this with the results of a small questionnaire in Dutch, and show that the generalisation holds for all persons (Section 3.1.2). Instead, plural should be analysed as '+ associates' (*a*) (as already suggested by a.o. Bobaljik (2008) and Ackema & Neeleman (To appear)), which is distinct from third person (*o*).

3.1.1. Morphology of plural and third person

If third person and plural were indeed the same, we would expect languages to exist with morphologically transparent paradigms (e.g. agglutinative languages) where plural and third person are expressed by the same morpheme, as in the following paradigm:[4]

(6) Hypothetical paradigm for 'plural = $+3^{rd}$ person':

PERS	SG	PL
iu		$\alpha(\text{-}\delta)$
i	β	$\beta\text{-}\delta$
u	γ	$\gamma\text{-}\delta$
o	δ	$\delta(\text{-}\delta)$

---

[4]The claim that third person and plural are the same is more often than not implicit in the literature, and is reflected in a type of notation that represents plural pronouns as '1+3', '2+3', etc. (e.g. Cysouw 2003), or *io*, *uo*, etc. (e.g. Harbour 2016). This does not necessarily mean that the relevant authors are committed to the belief that the third person and plural are identical. For this reason, I focus on Forchheimer (1953), since he has made the claim in the most explicit form that I am aware of. His claim is moreover formulated at a very general level, and ignores the distinction between semantic atoms (referents) and morphological atoms (features). My argument against the claim is based on a version that would assume that not only featural but also semantic identity will be reflected in a formal identity of some sort, and explores the consequences of such an assumption.

However, in the 370 languages from the sample, this pattern does not occur.

The only defense of the claim that 'plural = $+3^{rd}$ person' in terms of the morphology of attested paradigms is Forchheimer (1953), in which some languages are described as having the third person marker as a plural marker. However, for each of these languages, a more correct reanalysis shows that these are not true examples of 'plural = $+3^{rd}$ person' languages. I demonstrate this by discussing one of these composite forms, the independent pronouns of Kalaw Lagaw Ya (Pama-Nyungan). Similar reanalyses are available for the other languages, Arrernte (Pama-Nyungan), Siuslaw, Coos and Chinook (Penutian) and Hurrian and Sumerian; however, discussing all of them is beyond the scope of this paper.

Forchheimer mentions the following links between third person and plural in Kalaw Lagaw Ya (KLY) (Forchheimer 1953:127-128):

(7)　　a.　The exclusive plural *ngoi* 'may be a combination of [first person singular] *ngai*… and [third person singular] *noi*' (Forchheimer 1953:127).
　　　　b.　*Tana*, third person plural, is used as a plural suffix in the second person *ngitana* (Forchheimer 1953:127-128).

The pronoun paradigm given by Forchheimer is represented in (8) with the pronouns under discussion boldfaced.

(8)　　KLY personal pronouns by Forchheimer (1953:127)

	SG	PL
iu		*ngalpa*
i	**ngai** / *nazo*	**ngoi**, *ngöi*
u	*ngi*	*ngi***ta(na)**
o	**noi**, *nu* / *na, nadu*	**tana**, *ari*

If exclusive were a clear example of a '+ 3rd person plural' as suggested in (7a), we would expect to see *ngainoi* as the plural of first person, with the third person singular *noi* added to the first person singular *ngai*. This is not the case, so the link suggested by Forchheimer is more subtle: he seems to suggest that a part of the third person singular, *-oi*, is added to a part of the first person singular, *ng-*. Below, I show how other, more recent, analyses of KLY show no such link between the exclusive and third person singular, and that second person plural is formed by adding a distinct plural marker to second person singular, rather than adding third person plural to second person singular as suggested in (7b).

In Ray's description of KLY in (9) (1907:22), both the exclusive and the second person plural have slightly different forms. The same goes for the descriptions of Ford & Ober (2004:138) and Round & Stirling (2015:264), shown in (10).

(9)　　KLY pronouns by Ray (1907:22)

	SG	PL
iu		*ngalpa*
i	*ngai*	**ngoi**
u	*ngi*	**ngita**
o	**nui** / *na*	**tana**

(10)　　KLY pronouns by Ford & Ober (2004:138) and Round & Stirling (2015:264)

	SG	PL
iu		*ngalpa*
i	*ngay*	**ngoey**
u	*ngi*	**ngitha**
o	**nuy** / *na*	**thana**

The link Forchheimer mentions between exclusive and third person singular (7a) does not hold. As opposed to Forchheimer's *noi* for third singular, Ray notes *nui* and Ford & Ober and Round & Stirling note *nuy* as third person singular. Their exclusive *ngoi* and *ngoey* do not show a clear relation to third singular.[5]

As for (7b), the second person plural pronoun in (9) and (10) lacks the optional *-na* that Forchheimer analyses as a part of the third person plural pronoun *tana*. As such, second person plural *ngita* is a combination of second person singular *ngi* and *-ta*, the latter which Ray analyses as a plural marker. Support for Ray's claim comes from the KLY deictic system discussed by Ford & Ober and shown in (11), where *-tha* (or in Ray's paradigm, *-ta*) does indeed occur as a plural marker. As such, *ngi-t(h)a* should be decomposed as 2sg+pl and *t(h)a-na* as pl+3sg.[6]

(11)   The KLY deictic system (with more distinctions in Ford & Ober 2004:137)

	SG MASC	SG FEM	DU	PL
proximate	*in*	*ina*	*ipal*	*i**tha**
remote	*senaw*	*sena*	*sepal*	*se**tha**
'over there'	*pinungap*	*pinangap*	*pipalngap*	*pi**tha**ngap*
'up there'	*pinuka*	*pinaka*	*pipalka*	*pi**tha**ka*
'down there'	*pinuguy*	*pinaguy*	*pipalguy*	*pi**tha**guy*
'up at the front'	*pinupay*	*pinapay*	*pipalpay*	*pi**tha**pay*
'down at the back'	*pinupun*	*pinapun*	*pipalpun*	*pi**tha**pun*

In Ray (1907); Ford & Ober (2004); Round & Stirling (2015), there is no morphological link between exclusive and third person, nor between second person plural and third person. Therefore, KLY is not a convincing example of a language that uses the third person morpheme to express plural on the other persons.

Forchheimer's alleged examples of 'plural = +3rd person' languages are far from convincing (1953). Taken together with the absence of the paradigm in (6) in the 370 languages I looked at, we can thus conclude that it is not a possible paradigm.

The next section explains how semantics and more specifically the reference of plural pronouns confirms this.

3.1.2.   *Semantics of plural and third person*

In this section, I show that plural pronouns cannot refer to a group that adds 'others' or third persons to the relevant person. As such, *io* and *uo* do not correspond to respectively first and second person plural, but are unlexicalisable combinations, as further argued for in Section 3.2.

Ackema & Neeleman (To appear) also argue for a distinction between plural and third person. They distinguish two atoms in the person ontology: associates (*a*), being the non-focal referents of a plural pronominal, and others (*o*), being the third person referent.[7] They impose the same restriction as discussed in this paper: the atoms *i* and *u* cannot occur with *o*. Instead, these are pluralised by adding *a*. However, they have to stipulate this, as they do not explain why

---

[5]How the exclusive is to be decomposed remains unclear.

[6]It is unclear why plural is a suffix on second person and a prefix on third person. However, it is not uncommon for third person to behave differently from the other persons.

[7]In fact, they introduce two different kinds of associates: a_i and a_u, associates to the speaker and associates to the hearer. However, nothing in their system relies on this distinction and a_i and a_u can easily be conflated into one kind of associate.

o cannot occur with *i* or *u*.[8] The analysis introduced in Section 5, on the other hand, shows this cooccurrence restriction to be part of a generalisation on concept formation that holds for lexicalisation in general. Also, they do not extend this restriction to third person, which they analyse as one *o* for third person singular and several *o*s for third person plural. However, I argue that plural is formed by adding associates for all persons.

The argument they use to distinguish *a* from *o* is of a semantic nature:

> [W]hat is an *o* at a particular point in the discourse cannot be included in the reference of a first or second person plural pronoun without first being turned into an associate in some way. The other side of the coin is that an associate cannot be treated as an *o* without first being introduced as such (Ackema & Neeleman To appear:71).

This is illustrated with the example in (12): answer a. is considered odd because at this point in the conversation, George Clooney and Ad have not been established as associates. In this context, answer b. would be better. If a. is indeed uttered, then this has the comical effect of implying that George Clooney and Ad are in fact close friends.

(12)   (Peter:) Do you know whether George Clooney likes good coffee?
    a.   (Ad:) #Yes, we both drink Illy.
    b.   (Ad:) Yes, he drinks Illy, just like me.

However, after uttering b., an associate relation is established, namely the one of drinking Illy. This means that from this point on in the conversation, Ad can use a plural pronoun to refer to himself and George Clooney. For example:

(13)   (Following to the context in (12b)) We know good coffee when we see it.

This suggests that a plural pronoun is not made up of '+ third person' but rather of '+ associates' and that third person and associates are distinct elements.

In order to check whether this effect corresponds to speakers' judgements, I have conducted a small survey including 32 speakers of Flemish Dutch. The informants had to rate the suitability of the answers in situations such as the one described in Ackema & Neeleman (To appear), shown above, on a scale of 0 to 5 with 0 being a very odd and unexpected answer and 5 being a perfectly suitable answer. The questions were randomised. I briefly discuss the results for the example in (12) in (14)-(16), and afterwards, show that the same effect holds for third person in (17)-(19).

For each person discussed, I provide examples with and without an associate relation, and with a plural pronoun or a combination of two singular ones. For ease of exposition, I have underlined the third person under discussion and boldfaced the pronoun that is meant to refer to this. The example sentences discussed here are given in English.

---

[8]The non-occurrence of *o* with either *i* or *u* follows from the structure of their input set, which is a set of nested structures on which the person features operate. *o* cannot occur with *i* or *u* because their nested structures cannot be selected by these features to derive plural pronouns. However, they provide no motivation as to why this input set is structured in this particular way.

(14)   **First person, no associate relation:**
         Peter and Ad are watching a movie starring George Clooney. Peter asks Ad: 'Do you
         know whether <u>George Clooney</u> likes to drink coffee?' to which Ad replies:

   a.   'Yes, **we** both like to drink Nespresso.'
   b.   'Yes, **he** likes to drink Nespresso, just like me.'

(15)   **First person, associate relation:**
         Peter and Ad are watching a movie starring George Clooney. Ad is telling Peter that
         he and George Clooney go for a walk together every Sunday and that they have been
         friends for years now. Peter asks Ad: 'Do you know whether <u>George Clooney</u> likes to
         drink coffee?' to which Ad replies:

   a.   'Yes, **we** both like to drink Nespresso.'
   b.   'Yes, **he** likes to drink Nespresso, just like me.'

The table in (16) shows the average score and standard deviation for each example, with the
score boldfaced for the distinctions relevant to the issue at hand.[9]

(16)   Judgements for first person plural

Example	Associate relation	Number of the pronoun	Response (0-5)	SD
(14a)	x	pl	**2.44**	1.58
(14b)	x	sg	4.18	0.95
(15a)	✓	pl	**3.38**	1.44
(15b)	✓	sg	4.41	0.74

The response in the sentence without an associate relation was rated considerably better for the
singular pronouns (14b) than for the plural pronoun (14a): 4.22 versus 2.45. This confirms that
a third person (i.e. George Clooney) cannot be included in the reference of a first person plural
pronoun. Interestingly, when establishing an associate relation as in (15), the response with the
plural pronoun is now rated significantly better (with a p-value of 0.001): 3.39 in (15a) for *we*
compared to 2.45 in (14a) without the associate relation. However, even in the sentence with
the associate relation, the first person plural pronoun is still rated less good than the singular
pronouns (with a p-value of < 0.001): 3.39 versus 4.42. This is because an associate relation is
not entirely fixed and can be perceived differently by different people to a certain degree.

According to Ackema & Neeleman, only a first and second person plural pronoun cannot
include reference to third person. However, a similar example for third person shows the same
effect.[10] See for example (17) without and (18) with a possible associate relation. In both cases
the plural pronoun is rated less good when it includes the third person (in this situation: Julia
Roberts) but it is considered better when that person is an associate, as the numbers in (19)
show.

---

[10]I am discussing the animate use of third person pronouns. Third person pronouns referring to inanimates such
as 'Have you seen <u>the spoons</u>? **They** are not in the drawer' have a homogenous plural just like the nouns they refer
to, and are not discussed in this paper.

(17) **Third person, no associate relation:**
Peter and Ad are watching a movie starring George Clooney and <u>Julia Roberts</u>. Peter asks Ad: 'Do you know whether George Clooney likes to drink coffee?' to which Ad replies:

a. 'Yes, **they** both like to drink Nespresso.'
b. 'Yes, he likes to drink Nespresso, just like **Julia Roberts**.'

(18) **Third person, associate relation:**
Peter and Ad are watching a movie starring George Clooney and Julia Roberts. Peter asks Ad: 'Don't you think <u>Julia Roberts</u> and George Clooney act so well together? …By the way, do you know whether George Clooney likes to drink coffee?' to which Ad replies:

a. 'Yes, **they** both like to drink Nespresso.'
b. 'Yes, he likes to drink Nespresso, just like **Julia Roberts.**"

The table below shows the scores for the different answers.

(19)   Judgements for third person plural

Example	Associate relation	Number of the pronoun	Response (0-5)	SD
(17a)	x	pl	**1.79**	1.29
(17b)	x	sg	4.15	1.18
(18a)	✓	pl	**2.31**	1.40
(18b)	✓	sg	3.62	1.33

Just like the numbers for first person have shown above, the score for the answer with the plural is significantly better with the associate relation (18a) than without (17a) (with a p-value of 0.020).[11] This confirms that third person plural includes reference to associates rather than to third persons just like first and second person do.

Even though the associate relation discussed by Ackema & Neeleman is subject to some personal variation (in the sense that what counts as an associate relation for some may not for others) the results of the survey confirm that an other (*o*, third person) cannot be included in the reference of a plural pronoun without first being turned into an associate. This goes for all person distinctions in Dutch. As such, plural is formed by adding associates and *io* and *uo* do not refer to a plural first or second person. This will be relevant in what follows, where I discuss the lexicalisations or lack thereof of all logically possible person atoms and combinations.

3.2. *Person atoms and their combinations*

Consider again the person atoms and their logically possible combinations:[12]

---

[11]The answer with the singular pronoun in the situation with the associate relation, (18b), is rated less good than the corresponding answer without the associate relation in (17b). At this point it is unclear why this is, although it may be related to the repetition of the proper name in (18a).

[12] I assume that the referent *a* is not a person atom, since it differs from *i, u* and *o* in crucial ways. For example, it can never be spelled out on its own, and it always takes on the characteristics of whatever it occurs with. This is not the case for for example *u* in *iu*: it can be spelled out on its own, as second person, and it does not take on any characteristics of the *i* it occurs with in the inclusive. Ackema & Neeleman (To appear) argue that *a* is not a

(20) a. *i*   d. *iu*
      b. *u*   e. *io*
      c. *o*   f. *uo*

The single atoms of person can be referred to by first, second and third person. The combination *iu* is referred to by the inclusive. If *io* and *uo* could be lexicalised in a similar way to inclusive, we would expect them to show the same paradigmatic distinctions the other persons do.

Consider the three-way singular-plural and minimal-augmented paradigms shown below. As explained in the introduction, singular-plural languages make a number division based on cardinality (i.e. one atom for singular, two for dual and more than two for plural) (21). Minimal-augmented languages on the other hand, make a number distinction based on presence or absence of associates (i.e. no associates for minimal, + one associate for unit-augmented and + more than one associates for augmented) (22). If *io* and *uo* were lexicalisable person distinctions, we would expect a 15-cell paradigm for singular-plural languages: 3 single atoms × 3 numbers + 3 combinations × 2 numbers (i.e. 3 persons for singular, 6 for dual and 6 for plural). For minimal-augmented languages, we would expect an 18-cell paradigm: 6 persons × 3 numbers (i.e. 6 for minimal, 6 for unit-augmented and 6 for augmented).

(21) Singular-plural

SG	DU	PL
	io	*ioaaa...*
	uo	*uoaaa...*
	iu	*iuaaa...*
i	*ia*	*iaaa...*
u	*ua*	*uaaa...*
o	*oa*	*oaaa...*

(22) Minimal-augmented

MIN	UNIT	AUGM
io	*ioa*	*ioaaa...*
uo	*uoa*	*uoaaa...*
iu	*iua*	*iuaaa...*
i	*ia*	*iaaa...*
u	*ua*	*uaaa...*
o	*oa*	*oaaa...*

These paradigms are not attested. The maximally distinct paradigm that is attested for singular-dual-plural, has 11 cells: 3 single atoms × 3 numbers + 1 combination × 2 numbers (i.e. 3 for singular, 4 for dual and 4 for plural). An example of this are Tümpisa Shoshone pronouns in (23). In distributing the personal pronouns over the paradigm, it is clear that there are no pronouns that could express the combinations *io* and *uo*. None of the existing forms can refer to a combination of hearer and other (*uo*) or of speaker and other (*io*). (Remember that e.g. *mümmü*, the second person plural, cannot refer to *uo* but refers to *ua* instead, as argued for in the previous section.)

---

number atom either, for which they make a convincing argument. However, they draw the conclusion that it must then be a part of the person system. Based on a.o. the distinctions between *a* and *i, u, o* just mentioned, it is unclear why they would assume this. Where exactly the referent *a* does enter the person and number derivation, is left for future research.

(23)   Tümpisa Shoshone (Dayley 1979)

SG		DU		PL	
		io	*	ioaaa...	*
		uo	*	uoaaa...	*
		iu	*ta-ngku*	iuaaa...	*ta-mmü*
i	*nü*	ia	*nu-ngku*	iaaa...	*nü-mmü*
u	*ü*	ua	*mu-nku*	uaaa...	*mü-mmü*
o	(demonstr)	oa	(demonstr)	oaaa...	(demonstr)

The maximally distinct paradigm attested for minimal-unit-augmented has 12 cells: 4 persons × 3 numbers (i.e. 4 for minimal, 4 for unit-augmented and 4 for augmented). These distribute over the paradigm as shown in (24) for Rembarrnga. There are pronouns for the four person distinctions *i, u, o, iu* in all three numbers but not for *io* and *uo* in any number.

(24)   Rembarrnga (Cysouw 2003)

MINIMAL		UNIT-AUGMENTED		AUGMENTED	
io	*	ioa	*	ioaaa...	*
uo	*	uoa	*	uoaaa...	*
iu	*yʉkkʉ*	iua	*ngakorrʉ*	iuaaa...	*ngakorrbbarrah*
i	*ngʉnʉ*	a	*yarrʉ*	iaaa...	*yarrbbarrah*
u	*kʉ*	ua	*nakorrʉ*	uaaa...	*nakorrbarrah*
o	*nawʉ/ ngadʉ*	oa	*barrʉ*	oaaa...	*barrbarrah*

To conclude: the sample does not have a single instance of a lexicalised *io* and *uo*. In the next sections (Sections 4-5), I show that this gap in lexicalisation is part of a broader generalisation on limitations on concept formation. Before that, I shortly explain the difference between a lexicalised combination and a syncretism for two separate person referents, as at first sight, syncretisms between *i* or *u* and *o* may look like lexicalisations of the combinations *io* and *uo*. I show that they are not.

3.3.   Syncretism between u *and* o *or* i *and* o

Some languages have a single pronoun to refer to both second and third person, or to both first and third person. I show that these are not instantiations of a lexicalised *uo* or *io* combination, since those would refer to the mereological sum of both atoms, whereas a syncretism refers to a disjunction.[13] An example is the pronoun for second and third person singular in Sanapaná, *hlejap*:

---

[13]If the pronouns are considered as accidentally homophonous instead of syncretic, there are two separate lexical entries that each refer to just one corner of the kite. As such, they do not form a problem for the kite to begin with. However, I follow analyses like Distributed Morphology (Halle & Marantz 1993) and Nanosyntax (Caha 2009) in that these are in facts syncretisms and therefore correspond to a single lexical entry that can be inserted in two syntactic structures. In this case, one form corresponds to the disjunction of two referents, as argued for in this section.

(25)   Sanapaná (Harbour 2015)

	SG	PL
iu		*enenko'o*
i	*ko'o*	*enenko'o*
u	*hlejap*	*hlengap*
o	*hlejap*	*hlengap*

The interpretation of inclusive shows that lexicalised person combinations are mereological sums (I return to this distinction in Section 5.1). This means that the inclusive refers to the sum of speaker and hearer, *i* and *u*, but not to their disjunction, *i* or *u*. Consider again Tümpisa Shoshone, which has an established lexicalised combination: *tammü* for *iu*. Crucially, this pronoun can only refer to the mereological sum of speaker and hearer, not to either the speaker or the hearer, as the example below shows. If reference to either the speaker or the hearer is needed, a first or second person pronoun is used.

(26)   a.   **Tammü** piiya hipi"- tükin- tu'ih
             we.incl beer drink- start- will
             '**We**'ll start to drink beer.'   (Dayley 1979:14)
        b.   **Nü** tattanguhi punikka
             I   men.DU.O see
             '**I** see two men.'   (Dayley 1979:224)
        c.   **Mümmü** ha namo'okoitsoimmaa?
             you-PL   Q handwashed
             'Did **you** wash your hands?'   (Dayley 1979:325)

In contrast, syncretic pronouns are not mereological sums but disjunctions instead. They are merely pronouns that can be used for more than one concept, as in non-inclusive languages like English. *We* can be used to refer to the inclusive person or to the first person plural, depending on the context. Consider the examples below, where *we* in (27a) refers to the speaker and the hearer (and possibly associates) *iu(a)*, but *we* in (27b) does not refer to the hearer, therefore corresponding to *ia* (i.e. the speaker and his associates). This shows that in syncretism, there is no obligatory mereological sum, as opposed to lexicalised combinations.

(27)   Context: You are talking to someone from a different university whom you just met at a conference:
        a.   Don't you think **we** get excellent coffee at this conference?
        b.   In my office in Brussels, **we** have excellent coffee.

Now consider languages like Sanapaná in (25), where one pronoun, *hlejap* (and *hlengap* in plural), can refer to *u* or to *o*. *Hlejap* and *hlengap* could mistakenly be analysed as lexicalised *uo* combinations; however, I show that they are not. They behave like the English *we* rather than the Tümpisa Shoshone *tammü*, and are therefore not lexicalised combinations but rather syncretisms. The Sanapaná sentences below show that the pronoun *hlejap* refers to either second or third person, depending on the context, but never to the sum of both, as a plural translation 'you and him' for *hlejap* is simply not found in the data. As such, this is not the mereological sum *uo* and therefore not a lexicalised combination.

(28)   a.   **Hlejap** metko patakon ap-   angok.
                2sg/3sg NEG   money 2sg/3sg- POSS
                '**He** doesn't have money.'

        b.   Ta'asak akjehlna ap-   ta= o **hlejap**?
                which fruit   2sg/3sg- eat- Q **2sg/3sg**
                'Which fruit did **you** eat?'   (Harbour 2016:62)

The same holds for languages with a first and third person syncretism. These are not lexicalised combinations, but rather syncretisms of two different referents.
Consider for example the Spanish imperfect person suffixes:[14]

(29)   a.   habla- ba-   **ø**
                speak- PAST- **1/3SG**
                '**I/she/he** spoke'

        b.   habla- ba-   **s**
                speak- PAST- **2SG**
                '**You** spoke'   (Cysouw 2003:46)

Since Spanish does not require pronouns to be present, the sentence in (29a) is ambiguous: it could refer either to a first person or to a third person, but necessarily not to a combination of both, as I am not aware of such a (necessarily plural) interpretation. As such, this is not an instance of a lexicalised *io* combination.

The data from the sample demonstrates that from all logically possible person atom combinations, *io* and *uo* cannot be lexicalised in natural language. In what follows, I show how this limitation on the lexicalisation of person is part of a broader generalisation on concept formation in closed lexical fields, as observed in the kite framework. First, I discuss the framework itself (Section 4) to then explain how it applies to person and predicts exactly the pattern I have just shown (Section 5).

4.   *The kite framework*

In this section, I briefly explain the basis of the kite framework, and exemplify it with the lexical field of quantifiers. Then, I discuss the prediction this framework makes about lexicalisation in closed lexical fields: the CONCEPT FORMATION CONSTRAINT (Jaspers 2012; Seuren & Jaspers 2014; Roelandt 2016).

4.1.   *The logical square and hexagon*

The kite framework represents the concepts of closed lexical fields and the logical relations that hold between them in geometrical figures (Seuren & Jaspers 2014). This is a tradition that goes back to Aristotle's square of opposition, shown in Fig. 2. The square can be split up into two

---

[14]I am not aware of pronouns that show this syncretism in the singular as it is an uncommon pattern. Cysouw (2003) does list instances of plural pronouns with this syncretism, e.g. Bagirmi. The singular examples illustrate the point made more clearly, even though the plural syncretism shows the same kind of disjunction as the singular does. Therefore, this example is one of agreement markers rather than independent pronouns.

subuniverses: a positive one, with a universal affirmative (A) and a particular affirmative (I) (for the 'A' and 'I' in *affirmo*); and a negative one, with a universal negative (E) and a particular negative (O) (for the 'E' and 'O' in *nego*). The square can model closed lexical fields such as the quantifiers in Fig. 3.
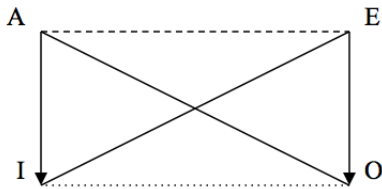

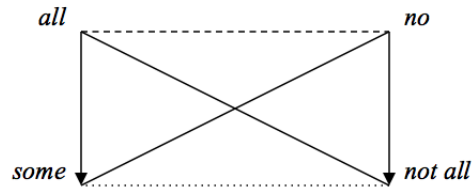*Figure 2*: *Aristotle's square of opposition*

Figure 3: *The square of opposition: quantifiers*

Between the corners, certain logical relations hold (Jaspers 2012):

(30)   a.   Entailment (arrows):
           (i)   If A is true, then I is also true.
           (ii)   If E is true, then O is also true.
           (iii) E.g.: If *all* Belgians love chocolate, then *some* Belgians love chocolate.
       b.   Contradiction (full lines):
           (i)   A and O cannot both be true or both be false at the same time.
           (ii)   I and E cannot both be true or both be false at the same time.
           (iii) E.g. '*All* Belgians love chocolate' and '*No* Belgians love chocolate' cannot be true or false at the same time.
       c.   Contrariety (dashed line):
           (i)   A and E cannot both be true but they can both be false at the same time.
           (ii)   E.g. '*All* Belgians love chocolate' and '*No* Belgians love chocolate' cannot be true at the same time, but they can be false at the same time, i.e. when only *some* Belgians love chocolate.
       d.   Subcontrariety (dotted line):
           (i)   I and O can both be true but not both be false at the same time.
           (ii)   E.g. '*Some* Belgians love chocolate' and '*Not all* Belgians love chocolate' can both be true at the same time, but they cannot be false at the same time.

It has long been noted that the quantifier *some* can be ambiguous. Consider the following sentences:

(31)   If some students pass the test, I'll treat them with chocolates.

(32)   Some students have passed the test.

Some in (31) can also refer to *all*, since I will still treat the students to chocolate if all of them pass the test. It is therefore also called the 'inclusive *some*', since it can include *all* (Roelandt 2016:108). If all students passed the test, the sentence in (32) would be false, since here *some* does not mean *all*. This is called the 'exclusive *some*'.

Jacoby (1950, 1960), Sesmat (1951) and Blanché (1952) attribute this difference to the existence of two different quantifiers *some*: an inclusive and an exclusive one (Jaspers 2012; Seuren & Jaspers 2014). Therefore, they extend the square to a hexagon (Figs. 4 and 5), which now

includes both kinds of *some* (in I and Y) and contradictories to each of them (in E and U respectively).[15]

Nine more relations run to and from these new corners: a contradiction between Y and U, four extra entailment relations, a triangle of contrariety (A-Y-E) and a triangle of subcontrariety (I-U-O). The triangle of contrariety is made up of concepts that consist of the conjunction of their adjacent corners (e.g. *some* in the Y-corner: '*some* and *not all*') and the triangle of subcontrariety is made up of corners that consist of the disjunction of their adjacent corners (e.g. the I-corner: '*all* or *some$_{excl}$*').
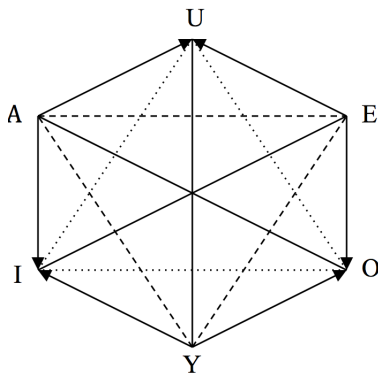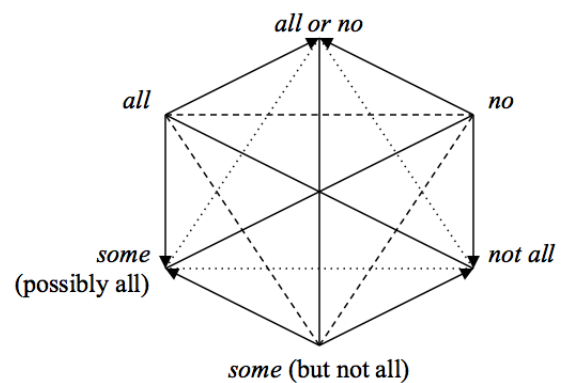

Figure 4: *Hexagon*

Figure 5: *Quantifier hexagon*

This hexagon now shows all the distinctions in the conceptual field of quantifiers and the logical relations that hold between the concepts.

4.2. The Concept Formation Constraint

Jaspers (2005, 2012) and Seuren & Jaspers (2014) note that not all of the corners on the hexagons can be lexicalised: both the O- and the U-corner do not receive a simplex lexicalisation in natural language.[16] This is generalised in the CONCEPT FORMATION CONTRAINT.

(33)   THE CONCEPT FORMATION CONSTRAINT (CFC)
        When mapping a closed lexical field onto the logical hexagon, the O- and U-corners never receive a simplex lexicalisation.

For quantifiers, this means that we cannot express *not all* (O) and *all or no* (U) as simplex lexical items. Omitting these O- and U-corners from the hexagons in Figs. 4 and 5 results in the kite.

---

[15]Grice (1975) attributes the two possible readings of *some* to pragmatic implicature, rather than to a semantic difference. In contrast, the approach in terms of the hexagon developed by Jacoby, Sesmat and Blanché takes the distinction between inclusive and exclusive *some* to be a semantic one (see Seuren & Jaspers 2014:620 for their arguments). Since this issue is orthogonal to my concerns, I do not discuss it any further here.

[16]For the O-corner, this was already observed by Thomas Aquinas and later generalised by Blanché (1953, 1966).
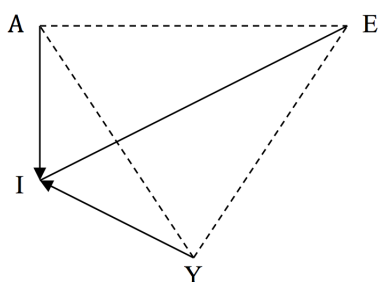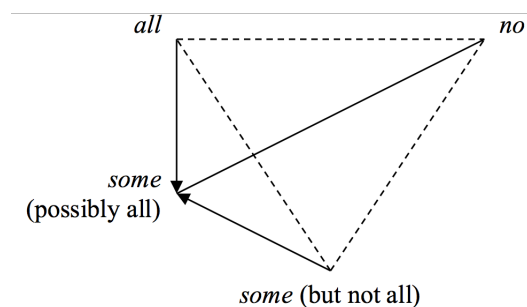
Figure 6: *Kite*

Figure 7: *Quantifier Kite*

The kite accounts for all the lexicalisable distinctions in the sense that it shows all the distinctions attested within the conceptual field of quantifiers and correctly rules out the unlexicalisable concepts. There is no simplex lexical item *allno* to express the disjunction *all or no*, neither is there a *nall* in natural language to express *not all*.

The CFC has been shown to apply to quantifiers as demonstrated above, to the predicate calculus logical operators *and, or, nor* and to binary lexical fields with an overarching hyperonym that can be split up into subdomains, for example 'human' as an overarching lexical predicate for 'man' and 'woman' by Seuren & Jaspers (2014). It has also been applied to colour perception terms in Jaspers 2012 and extended for adjectives by Roelandt (2016). In all these cases, when mapping the lexical field onto the hexagon, the same two corners do not receive a simplex lexicalisation. Details can be found in the works cited.

In what follows, I argue that this same CONCEPT FORMATION CONSTRAINT holds for person (as already hinted at in Seuren & Jaspers 2014) and show that the absence of a person morpheme for *io* and *uo* is part of this broader generalisation on lexicalisation limitations.

5.   *The person kite*

In this section, I explain how the logical relations mentioned above apply to person in Section 5.1. Section 5.2 presents how combining these relations with the person atoms results in the person kite and how the person concepts end up in their respective corners. Applying the CFC to this derived kite, will correctly predict the inclusive *iu* to be the only lexicalisable person atom combination since *io* and *uo* will be ruled out as lexicalisable in natural language.

5.1.   *The mereological relations of the person kite*

As explained in Section 4.1, the hexagon has relations of entailment, contradiction, contrariety and subcontrariety between its vertices. These are logical relations that hold between propositions. For quantifiers, for example, the entailment relation *all* ⊆ *some* really holds between propositions of the sort *All Belgians like chocolate* and *Some Belgians like chocolate*. Person atoms and their combinations, on the other hand, do not correspond to propositions but to person referents. Therefore, it is unclear how the logical relations described above could apply. For example, for an entailment to hold between *i* and *iu*, these person atoms would need to be propositions: e.g. 'x is a speaker' for *i*. However, for *iu* this would correspond with 'x is

a participant'. Under such an interpretation, there is an entailment from *i* to *iu*, since being a speaker entails being a participant. However, in this case *iu* would correspond to *i or u* (since being a participant does not mean being both a speaker and a hearer, but being either a speaker or a hearer) and that does not correspond to the reference of the inclusive person.[17] Instead, the inclusive refers necessarily to the mereological sum of the speaker and the hearer. Therefore, for person, I take the relations to be of a mereological nature.

Mereology is the theory of parthood relations (Varzi 2016). This means that rather than dealing with entailment relations between logical concepts and propositional truth values, a mereology deals with parthood relations between atoms in a universe and their mereological sums (such as cake being the mereological sum of flour, butter, eggs and sugar). For example, for person, both *i* and *u* are parts of *iu* and conversely, *iu* is the mereological sum of *i* and *u* ($i \oplus u$).

The relations in logical and mereological hexagons show a very clear isomorphism. The corresponding relations (and the labels used for both in this paper) are shown in the table in (34). (35) shows the logical relations already exemplified in Section 4.1 and (36) explains the corresponding mereological relations.

(34)   The corresponding logical and mereological relations

Logical	Mereological	Label
Entailment	Parthood	$\subseteq$
Contradiction	Exhaustive complementarity	CD
Contrariety	Non-exhaustive complementarity	C
Subcontrariety	Exhaustive overlap	SC

(35)   Logical relations:

    a.   **Entailment**: A entails B iff when A is true, B is also true.

    b.   **Contradiction**: A and B are contradictory iff they cannot both be true or both be false at the same time.

    c.   **Contrariety**: A and B have a contrariety relation iff they cannot both be true but they can both be false at the same time.

    d.   **Subcontrariety**: A and B have a subcontrariety relation iff they can both be true but not both be false at the same time.

(36)   Mereological relations:

    a.   **Parthood**: A $\subseteq$ B iff A is a part of B.

    b.   **Exhaustive complementarity**: A and B are exhaustive complements iff their intersection is empty (i.e. they have no overlapping elements) and their union equals the universe.

    c.   **Non-exhaustive complementarity**: A and B are non-exhaustive complements iff their intersection is empty but their union does not equal universe.

    d.   **Exhaustive overlap**: A and B show exhaustive overlap iff their intersection is non-empty (i.e. they have at least one overlapping atom); their union equals the universe, and neither is part of the other.

---

[17]Another possible entailment is the following: 'we (*iu*) like coffee' entails 'I (*i*) like coffee'. This, however, is an entailment that does not hold with collective predicates, i.e. 'we gathered in the hall' does not entail 'I gathered in the hall'. (I would like to thank Guido Vanden Wyngaerd for pointing this out to me.) In short, entailments, however construed, are problematic for the inclusive, and need to be replaced by mereological relations.

The isomorphism between these relations becomes clear when employing bitstrings to represent the corners of the hexagon (Smessaert 2009; Jaspers 2012; Roelandt 2016). In general, any hexagon can be described as composed of three atoms (*i, u* and *o* for person and *all, some* and *no* for quantifiers) and their combinations. The number of possibilities equals $2^3 = 8$. Discounting 000 and 111, this yields the six corners of the hexagon. Therefore, the bitstrings are of length three with every position in the bitstring corresponding to exactly one atom: a 1-value indicates the presence of this atom and a 0-value the absence. Combining the three atoms leaves us with eight logical possibilities, of which I will not discuss the absence of all atoms, 000, nor the presence of all atoms, 111 (as motivated in Footnote 2).[18]

(37)   Logical possibilities with bitstrings

Bits	Quantifiers	Person
100	*all*	*i*
010	*some* ($some_{excl}$)	*u*
001	*no*	*o*
110	*all or some* ($some_{incl}$)	*iu*
101	*all or no*	*io*
011	*some or no* (*not all*)	*uo*

These bitstrings can be used to formalise the logical and mereological relations explained above, and show how they are parallel (Smessaert 2009). For each mereological relation I mention the equivalent logical relation and the abstract formula for both, with ϕ and ψ representing the concepts on the vertices of the hexagon. In these formulas, intersection ($\cap$) selects only the atoms (i.e. 1-values) both items have in common, whereas union ($\cup$) selects all the atoms (i.e. 1-values) present in either or both of the items. For example, $100 \cap 010 = 000$ and $100 \cup 010 = 110$. Underneath the general formulas, I show how this applies to quantifiers in a. and to person in b. I also show how both are generalised in the same way by means of the bitstrings in c.

(38)   Parthood and entailment ($\subseteq$):
$\phi \subseteq \psi$ iff $\phi \cap \psi = \phi$ and $\phi \cup \psi = \psi$

   a.   *all* $\subseteq some_{incl}$ because *all* $\cap some_{incl} = $ *all* and *all* $\cup some_{incl} = some_{incl}$
   b.   *i* $\subseteq$ *iu* because *i* $\cap$ *iu* = *i* and *i* $\cup$ *iu* = *iu*
   c. $100 \subseteq 110$ because $100 \cap 110 = 100$ and $100 \cup 110 = 110$

(39)   Exhaustive complementarity and contradiction (CD):
ϕ CD ψ iff $\phi \cap \psi = 000$ and $\phi \cup \psi = 111$

   a. $some_{incl}$ CD *no* because $some_{incl} \cap$ *no* $= 000$ and $some_{incl} \cup$ *no* $= 111$
   b.   *iu* CD *o* because *iu* $\cap$ *o* $= \varnothing$ and *iu* $\cup$ *o* = *iuo*
   c. 110 CD 001 because $110 \cap 001 = 000$ and $110 \cup 001 = 111$

(40)   Non-exhaustive complementarity and contrariety (C):
ϕ C ψ iff $\phi \cap \psi = 000$ and $\phi \cup \psi \neq 111$

   a.   *all* C *no* because *all* $\cap$ *no* $= 000$ and *all* $\cup$ *no* $\neq 111$

---

[18]Note that whether or not 111 and 000 receive simplex lexicalisation can differ from lexical field to lexical field. For colour for example, 000 corresponds to black and 111 to white. These concepts are included in the hexagon with a third dimension: both are points in the centre, one sticking out on top, the other at the bottom Jaspers (2012). This is likely also the case for person.

   b.   *i* C *o* because $i \cap o = \varnothing$ and $i \cup o \neq iuo$

   c.   100 C 001 because $100 \cap 001 = 000$ and $100 \cup 001 \neq 111$

(41)   Exhaustive overlap and subcontrariety (SC):
   ϕ SC ψ iff $\phi \cap \psi \neq 000$; $\phi \cup \psi = 111$; $\phi \nsubseteq \psi$; $\phi \nsupseteq \psi$

   a.   *some*$_{incl}$ SC *not all* because *some*$_{incl}$ $\cap$ *not all* $\neq 000$; *some*$_{incl}$ $\cup$ *not all* $= 111$; *some*$_{incl}$ $\nsubseteq$ *not all*; *some*$_{incl}$ $\nsupseteq$ *not all*

   b.   *iu* SC *uo* because *iu* $\cap$ *uo* $\neq \varnothing$; *iu* $\cup$ *uo* $=$ *iuo*; *iu* $\nsubseteq$ *uo*; *iu* $\nsupseteq$ *uo*

   c.   110 SC 011 because $110 \cap 011 \neq 000$; $110 \cup 011 = 111$; $110 \nsubseteq 011$; $110 \nsupseteq 011$

To illustrate, consider the hexagons below with the bitstrings added to the corners.[19] (I explain exactly how the person hexagon is derived in the next section.)
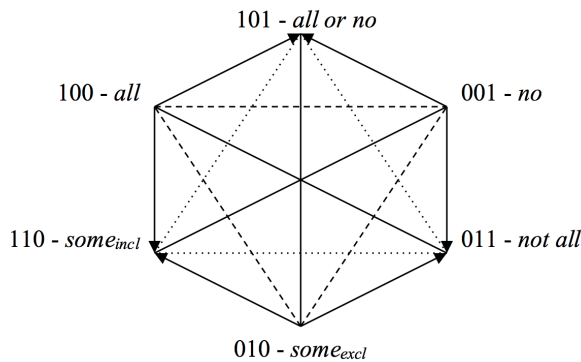

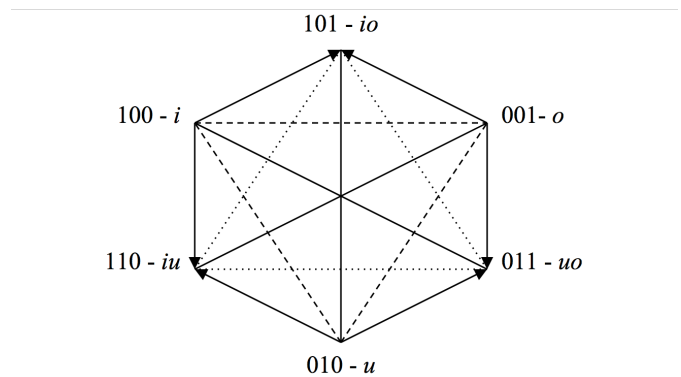Figure 8: *Quantifier hexagon with bitstrings*

Figure 9: *Person hexagon with bitstrings*

This shows that even though logical and mereological relations apply to different kinds of concepts (i.e. propositions vs. referents), they show an isomorphism, which can be brought out with the bitstring formalism.

    The crucial difference between logical and mereological lexical fields relates to the I-O-U corners (see Fig. 4 for the hexagon with letters representing the corners). In logical systems, these make up the triangle of subcontrariety and correspond to the disjunction of their adjacent corners. For example, for quantifiers, the I-corner *some*$_{incl}$ corresponds to *some*$_{excl}$ *or all* (Fig. 10). For mereologies, these corners (making up the triangle of exhaustive overlap) are the mereological sums ($\oplus$) of their adjacent corners (Fig. 11). This means that for the person hexagon I derive below, *iu* corresponds to *i*$\oplus$*u*.

---

[19]Note that which atom receives which 1-bit in the bitstring is arbitrary and does not affect the relations explained below. However, which concept ends up in which concept of the hexagon is not arbitrary, and will be explained in Section 5.
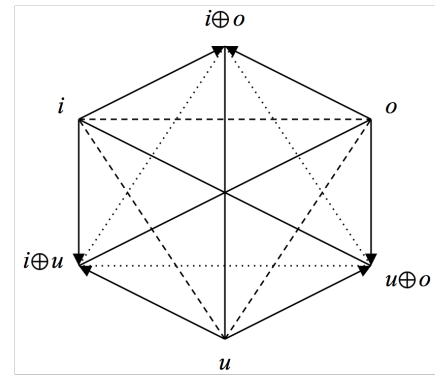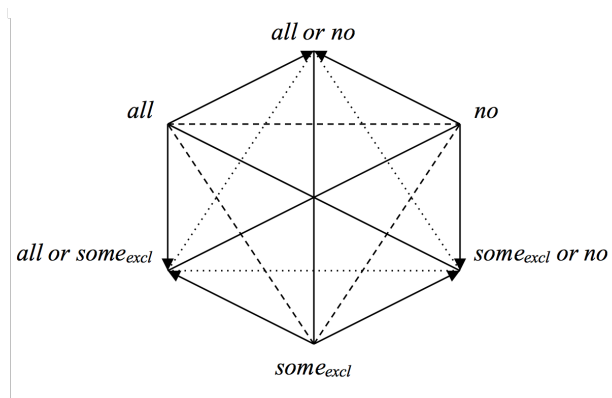
Figure 10: Quantifier hexagon: disjunction

Figure 11: Person hexagon: mereological sum

Despite this difference, mereological hexagons are also subject to the CFC, just like logical ones.[20]

5.2.   Deriving the person kite

Now that we have seen how the CFC applies to closed lexical fields and that person is a mereological lexical field, I show exactly how the person kite is derived and why the concepts end up in their respective corners.

Seuren & Jaspers (2014) show how from one single entailment relation between two concepts (e.g. *all* $\subseteq$ *some$_{incl}$*), a hexagon, and by virtue of the CFC also a kite, automatically follows. (To see how the parallel with any other lexical field works in detail, see Seuren & Jaspers 2014:612-613;626-627.) For a mereology, the hexagon is derived from a parthood relation: $i \subseteq iu$.

Here I explain the step by step derivation of the person kite and illustrate with Fig. 12.

- **Step 1**: The derivation of the person hexagon starts with the parthood relation between *i* and *iu*: *i* is a part of *iu* because it is contained in *iu*.

- **Step 2**: From any parthood relation, two exhaustive complementarity relations follow. One is between *iu* and *o*: *iu* and *o* are exhaustive complements because they have no overlapping elements and their union equals the universe (i.e. *iuo*). The other is between *i* and *uo*: *i* and *uo* also have no overlapping elements but their union equals the universe.

- **Step 3**: These two new corners stand in a parthood relationship: *o* is a part of *uo*. There is also a non-exhaustive complementarity relationship between the single atoms, *o* and *i*: they are non-exhaustive complements because they have no overlapping elements but their union does not equal the universe (i.e. *u* is still a possibility in the universe). Finally, there is an exhaustive overlap relation between the combinations *iu* and *uo*: they overlap exhaustively because their union equals the universe *iuo*, their intersection is non-empty because it contains *u*, and they are no subsets because *i* and *o* are complementary. The result is the mereological equivalent of the square of opposition.

---

[20]This has been demonstrated for the mereology of colour, see Jaspers 2012 for a full discussion.

- **Step 4**: The atoms *i* and *o* each have their own corner, but the atom *u* in *iu* and *uo* does not. Adding this final single atom, *u*, completes the universe and thereby finishes the triangle of non-exhaustive complementarity. There are also two more parthood relations: from *u* to *iu* and from *u* to *uo*.

- **Step 5**: There is now a new complementarity relation between *u* and *io*. There are also two more parthood relations, from *i* and *o* to *io*, and the triangle of exhaustive overlap is finished.

- **Step 6**: At this point, taking out the unlexicalisable corners *io* and *uo* turns the hexagon into the kite.
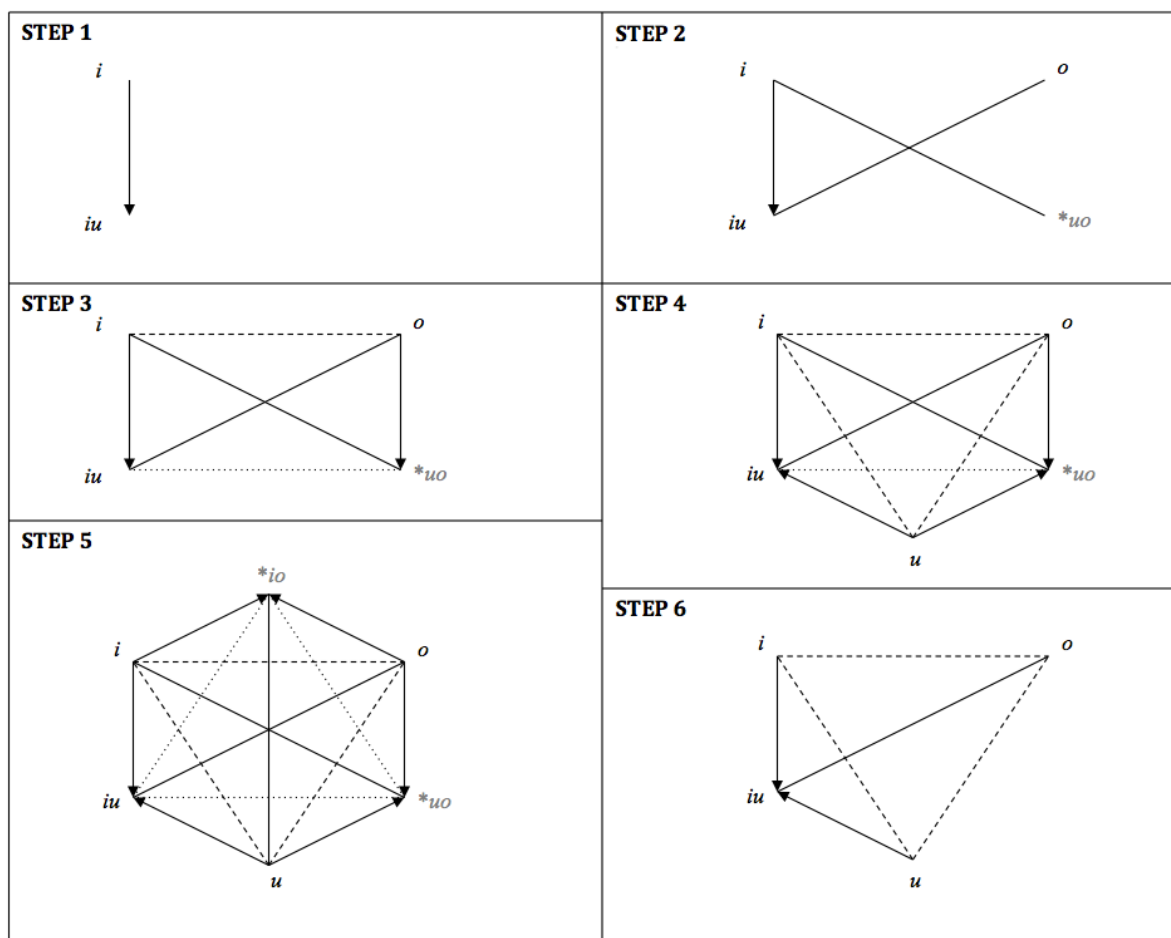
Figure 12: *The derivation of the person kite step by step*

Associates do not occupy a position in the hexagon, the reason for this being that they are not person atoms (as explained in Footnote 12).

Note that there are other parthood relations that can derive a hexagon, e.g. $o \subseteq io$. However, apart from representing the mereological relations, the hexagon also shows a divide between a positive universe (A-Y-I, remember: A and I in *affirmo* as explained in Section 4.1) and a negative one (U-E-O, E and O in *nego*) (Jaspers 2012:239; Seuren & Jaspers 2014:622). In the quantifiers for example, the only negative quantifier *no* resides both in the U and E corner, and in

the O-corner we find a contradiction of a positive quantifier. For colour, the colours in the positive universe are often considered as warm colours, whereas the ones in the negative universe are often considered as cold. Other parthood relations amongst person atoms, e.g. between *o* and *io*, may also provide a mereological hexagon, but this will not follow the positive-negative divide. Third person, *o*, belongs to the negative side of the person universe since, as opposed to the other person atoms, it does not refer to a speech act participant and, morphologically, it often differs from the other persons in for example selecting different number morphology or being syncretic with demonstratives (a.o. Silverstein 1976; Nevins 2007; Bobaljik 2008).

Therefore, the only two viable parthood relations to start from are *i* or *u* $\subseteq$ *iu*. Both result in a mereological hexagon with the participants on the positive side and non-participants on the negative side. So far, I have not come across any convincing arguments to start with one rather than the other. Note that this choice does not affect the predictions made for person lexicalisation: in both cases, *io* and *uo* will end up in the O- and U-corners and therefore be predicted to be unlexicalisable.

I illustrate this with the example from Section 3.2, Tümpisa Shoshone (23). The person morphemes available correspond to the corners of the kite and there are no morphemes for the concepts in the O- and U-corners (Fig. 13).

Note that the generalisation on the lexicalisation of person made in this paper concerns person morphemes and not full person and number markers. As such, the corners in the hexagons above represent only the person morphemes. For fusional paradigms like English (i.e. paradigms that do not have separate person and number morphemes but express both in just one morpheme) the kite looks more filled: the singular, plural, different genders, etc., are placed together in their relevant person corner. However, in fusional paradigms as well there is no pronoun to express a combination *io* or *uo* (Fig. 14).
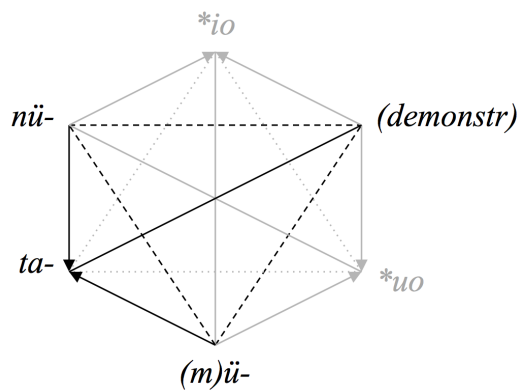
Figure 13: Hexagon for Tümpisa Shoshone
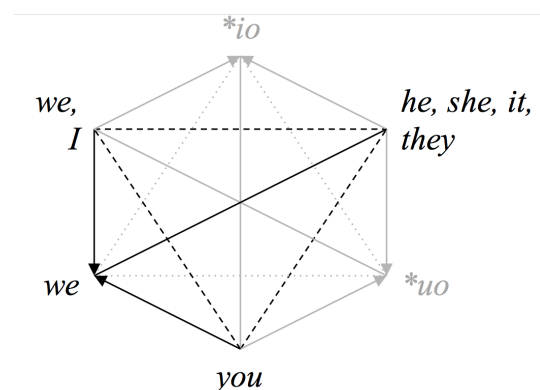
Figure 14: Hexagon for English

6.   Conclusion

I have shown that plural is not a case of '+ third person' since morphologically, no languages share a morpheme for third person and plural and semantically, the reference of a plural pronoun cannot include a third person (Ackema & Neeleman To appear). Instead, plural is a case of '+ associates'. In order for a third person to be included in the reference of a plural pronoun, it first needs to be established as an associate.

Also, syncretism between second and third person or first and third person is not an instantiation of a lexicalised person atom combination, since syncretic pronouns refer to only one of the persons it can express (depending on the situation), i.e. to $u \cup o$ and $i \cup o$, and not to the mereological sum of these, $u \oplus o$ and $i \oplus o$.

From all logically possible person referents and their combinations, two groups can never be lexicalised: *io* and *uo*. Exactly the gap in the lexicalisation of person that *io* and *uo* leave is predicted by the kite framework. Its CONCEPT FORMATION CONSTRAINT makes a prediction about lexicalisation in closed lexical fields: when representing the concepts of these lexical fields in a logical hexagon, the same two corners are never lexicalised in natural language. Applying this framework to person shows that the unlexicalisability of *io* and *uo* is part of this broader generalisation on limitations of concept formation.

Jolijn Sonnaert
KU Leuven - CRISSP
www.crissp.be
jolijn.sonnaert@kuleuven.be

Acknowledgements

References

Ackema, P. & A. Neeleman (To appear). *Features of Person*.

Baerman, M., D. Brown & G. G. Corbett (2005). *The Syntax-Morphology Interface*. Cambridge University Press, Cambridge.

Blanché, R. (1952). Quantity, modality and other kindred systems of categories. *Mind* 61:243, pp. 369–375.

Blanché, R. (1953). Sur l'opposition des concepts. *Theoria* :19, pp. 89–130.

Blanché, R. (1966). *Structures intellectuelles: Essai sur l'organisation systématique des concepts*. J. Vrin, Paris.

Bobaljik, J. D. (2008). Missing persons. *The Linguistic Review* 25, pp. 203–230.

Caha, P. (2009). *The Nanosyntax of Case*. Ph.D. thesis, University of Tromsø.

Cysouw, M. (2003). *The Paradigmatic Structure of Person Marking*. Oxford University Press, New York, 2009 edn.

Daniel, M. (2005). Understanding inclusives. Filimonova, E. (ed.), *Clusivity: Typology and case studies of the the inclusive–exclusive distinction*, John Benjamins Publishing Co., pp. 3–48.

Dayley, J. P. (1979). *Tümpisa (Panamint) Shoshone Grammar*. University of California Press, Berkely.

Dryer, M. S. & M. Haspelmath (eds.) (2013). *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Forchheimer, P. (1953). *The category of person in language*. Walter de Gruyter & Co, Berlin.

Ford, K. & D. Ober (2004). A sketch of Kalaw Kakaw Ya. Romaine, S. (ed.), *Language in Australia*, Cambridge University Press, Cambridge, chap. 7, pp. 118–142.

Grice, P. H. (1975). Logic and conversation. Cole, P. & J. L. Morgan (eds.), *Syntax and semantics*, Academic Press, New York, vol. 3, pp. 41–58.

Halle, M. & A. Marantz (1993). Distributed Morphology and the pieces of inflection. Hale, K. & S. J. Keyser (eds.), *The View from Building*, MIT Press, vol. 20, pp. 111–176.

Harbour, D. (2015). Poor pronoun systems and what they teach us. Krämer, M., S.-I. Ronai & P. Svenonius (eds.), *Nordlyd*, University of Tromsø, Tromsø, vol. 40.

Harbour, D. (2016). *Impossible Persons*. MIT Press, Cambridge, Massachusetts.

Harley, H. & E. Ritter (2002). Person and number in pronouns: A feature-geometric analysis. *Language* 78:3, pp. 482–526.

Jacoby, P. (1950). A triangle of opposites for types of propositions in aristotelian logic. *The New Scholasticism* :24, pp. 32–56.

Jacoby, P. (1960). Contrariety and the triangle of opposites in calid inferences. *The New Scholasticism* :34, pp. 141–169.

Jaspers, D. (2005). *Operators in the Lexicon*. Ph.D. thesis, Universiteit Leiden LOT, Utrecht.

Jaspers, D. (2012). Logic and colour. *Logica Universalis* 6, pp. 227–248.

Moskal, B. (2014). The role of morphological markedness in inclusive/exclusive pronouns. *Annual Meeting of the Berkeley Linguistics Society*, pp. 354–368, URL http://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/3148/2867.

Nevins, A. (2007). The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory* 25, pp. 273–313.

Ray, S. H. (1907). *Reports of the Cambridge anthropological expedition to Torres Straits*, vol. 3. Cambridge University Press, Cambridge.

Roelandt, K. (2016). Most *or the art of compositionality: Dutch* de/het meeste *at the syntax-semantics interface*. Ph.D. thesis, KU Leuven, LOT, Brussels.

Round, E. & L. Stirling (2015). Universals of split argument coding and morphological neutralization: Why Kala Lagaw Ya is not as bizarre as we thought. *Australian Journal of Linguistics* 35:3, pp. 251–281.

Sesmat, A. (1951). *Logique: Les Raisonnements. La Logistique.* Hermann.

Seuren, P. A. M. & D. Jaspers (2014). Logico-cognitive structure in the lexicon. *Language* 90:3, pp. 607–643.

Silverstein, M. (1976). *Hierarchy of features and ergativity*, Australian Institute of Aboriginal Studies, pp. 112–171.

Smessaert, H. (2009). On the 3D visualisation of logical relations. *Logica Universalis* 3, pp. 303–332.

Varzi, A. (2016). Mereology. *The Stanford Encyclopedia of Philosophy* URL http://plato.stanford.edu/archives/spr2016/entries/mereology/.

Phase, plurality and floating

Takanobu Nakamura

In this paper, I propose a modification of Kratzer's (2013) theory using a cyclic mapping between syntax and semantics. In the proposed account, the highest element within an extended projection of a lexical category is the phase edge and its sister constituent is transferred. So, the D's [plural] feature is always at the edge of the nominal phase and so interpreted as a *-operator within the verbal domain. In this approach, the syntactic theory is not complicated. Furthermore, the proposed system straightfowardly accounts for event-related readings of numerals in English and rigidly event-related nature of floating numeral quantifiers in Japanese.

1. Introduction

In this paper, I try to construct an event-based theory of plurality, incorporating phase-based cyclicity into compositional semantics. Specifically, phrasal plurality is obtained only when the [plural] feature is at the edge of the nominal phase (e.g., D/K). Comparing two major sources of plurality, Lexical Cumulativity and a covert distributive operator (a *-operator), Kratzer (2013) argues that Lexical Cumulativity alone undergenerates the relevant observations on distributive readings, whereas free insertion of a *-operator overgenerates them. Instead, she proposes a mixture of them, restricting *-operators to be associated with a verbal phrase only when it is the sister of a plural DP. To do this, she argues that D's [plural] feature is interpreted as a *-operator, moving out its DP and becoming verbal inflection. However, I argue that this treatment calls for several theoretical problems. To avoid them, while sustaining its empirical coverage, I propose that the D's [plural] feature is always interpreted within v/VP because the highest element within an extended projection is the phase edge and the edge of DP is not within the transfer domain of DP. In this approach, the problems mentioned above do not arise. The [plural] feature of D is interpreted as a *-operator for v/VP because it is always at the DP edge and interpreted within the v/VP. Moreover, the proposed account is superior to the previous one in describing event-related readings of numerals in English and rigidly event-related nature of floating numeral quantifiers in Japanese. The rest of this paper is organized as follows. In section 2, I review Kratzer's (2013) theory and argument for the mixture of lexical cumulativity and *-operators. In section 3, I point out that Kratzer's (2013) mechanism will lead to an undesirable complication of syntactic theory and I propose a phase-based modification to her model. Section 4 is for a discussion of floating numeral quantifiers in Japanese which is best captured by the proposed phase-based account.

2. Verbal plurality: the sources of plurality and Neo-Davidsonian notions of events

In this section, I briefly summarize the basic notions in theories of plurals and events, and Kratzer's (2013) argument that the mixture of Lexical Cumulativity and restricted use of *-operators best predict the availability of distributive readings. In section 2.1, I briefly sketch basic notions of the semantics of plurals. In section 2.2, I introduce two major sources of plurality: *Lexical Cumulativity* and a **-operator*. In section 2.3, I review Kratzer's (2013) event-based theory of plurals. Section 2.4. concludes this section.

2.1. Plurals as individual sums

First, plurals in natural language are treated as *individual-sums* (*i-sums*), e.g., 'John and Mary' denotes a *plural individual* which has John and Mary as its parts, but is distinct from John and Mary. To formally describe plural individuals, the i-sum forming operator '+' is defined (Link 1983). This operator has the following properties (Krifka 1998).

(1) a. $x+y = y+x$ (Commutativity)
    b. $x+(y+z) = (x+y)+z$ (Associativity)
    c. $x+x = x$ (Idempotence)

Individuals which has no parts are called *atomic individuals* and plural individuals are those which have other individuals as its parts. To express a part-whole relation between individuals, a relation '⊑' is defined, e.g., (2) expresses that John is a part of John and Mary.

(2) John ⊑ John+Mary

This operator '⊑' satisfies the following conditions (Krifka 1998).

(3) a. $x ⊑ x$ (reflexive)
    b. $(x ⊑ y \land y ⊑ x) \rightarrow x = y$ (antisymmetric)
    c. $(x ⊑ y \land y ⊑ z) \rightarrow x ⊑ z$ (transitive)

The domain of individuals must be closed under '+' to include both atomic individuals and plural individuals. So, whenever there are x and y in the domain of individuals D_e, so is $x+y$. The notion of plurals as i-sum well suits the semantics of collective predicates such as 'meet', 'gather' and so on because we can simply say that collective predicates are only true of plural individuals. However, if so, the notion of plurals as i-sum has a difficulty in describing distributivity, i.e., that a property is distributed to each atomic part of a plural individuals. To accommodate with distributivity, I use a **-operator*, which decompose a plural individual into atomic individuals which are the parts of it (cf. Link's 1983 distributive operator).

(4) $[[*]] = \lambda P_{<e, t>}. \lambda x. \forall y ⊑ x. [\text{Atom } (y) \leftrightarrow P(y)]$

'Atom (x)' is a meta-predicate which is true of an individual if and only if the individual is atomic. So, distributivity arises only when a *-operator appears and it decomposes a plural individual into atomic individuals. Next, I move on to the treatment of plural nouns. Following

Sauerland et al. (2005), I assume that denotation of plurals includes atomic individuals and denotations of plural nouns are as in (5).

(5) [[books]] = $\lambda x.\ \forall y \sqsubseteq x.$ [Atom $(y) \leftrightarrow$ book(y)]

Based on this, the plural morpheme '-s' is considered to have the following denotation.

(6) [[-s]] = $\lambda P_{<e,\ t>}.\ \lambda x.\ \forall y \sqsubseteq x.$ [Atom $(y) \leftrightarrow P(y)$]

Here, the plural morpheme '-s' is defined the same as *-operators. Lastly, I assume numerals to be modifiers of type $<e,\ t>$ as in (7).

(7) [[three]] = $\lambda x\ [|x| = 3]$

Now, we are ready to discuss pluralization mechanisms to account for the proper distribution of distributive reading. From the next section, I will compare *-operators with the other source of pluralization: *Lexical Cumulativity*.

2.2.   Lexical cumulativity and *-operators

In this subsection, I introduce *Lexical Cumulativity* and compare it with *-operators. First, *-operators are generalized to two-place predicates. When observing two-place predicates, we find a reading which is neither a collective reading nor a distributive reading. The reading is observed when a sentence has more than one plural DPs, e.g., (8a) allows both (8b) and (8c)

(8) a. The three boys invited the five girls.
    b. Each of the three boys invited each of the five girls.
    c. Each of the three boys invited at least one of the girls, and each of the five girls was invited by at least one of the boys.

The reading in (8c) has a distinct truth condition from that of (8b), e.g., (8a) is true in the following situation under the reading (8c), whereas false under the reading (8b).

(9)
$$
\begin{array}{ll}
 & g_1 \\
b_1 & g_2 \\
 & g_3 \\
b_2 & g_4 \\
b_3 & g_5
\end{array}
$$
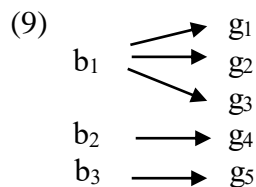
This reading is called a *cumulative reading*. To derive it, *-operators are generalized to a relation of type $<e,\ <e,\ t>>$ as follows.

(10) *R$(x)(y) = 1$ iff $x = x_1+x_2+\ \ldots\ +x_n$ and $y = y_1+y_2+\ \ldots\ +y_n$ and R$(x_1)(y_1) =$ R$(x_2)(y_2) = \ldots =$ R$(x_n)(y_n) = 1$

This generalized notion of decomposition of a plural individual into atomic individuals plays a crucial role in *Lexical Cumulativity*, which ascribes plurality to a lexically default option.

(11) <u>Cumulativity Universal</u>: Every simple predicate in natural language satisfies the cumulativity condition. (Krifka 1998 and Kratzer 2013 among others).

The cumulativity condition is described as follows.

(12) <u>Cumulativity Condition</u>: For any entities $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$, if $P(x_1) \ldots (x_n) = 1$ and $P(y_1) \ldots (y_n) = 1$, then $P(x_1+y_1) \ldots (x_n+y_n) = 1$ (Cable 2014)

If this hypothesis is correct, we have no need to posit a *-operator. In the next subsection, however, I will review Kratzer's (2013) argument that neither Lexical Cumulativity alone nor generalized *-operators makes right prediction and a mixture of them best analyzes the distribution of distributive/cumulative readings.

2.3.   Kratzer (2013): A mixture of Lexical Cumulativity and restricted *-operators

In this subsection, I review Kratzer's (2013) mechanism which makes use of both Lexical Cumulativity and *-operators. In her system, all predicative stems are born as plural and in addition, DPs with [plural] feature can release this feature to pluralize adjacent verbal projections. There is one more essence in her system; the notion of events. In non-event-based semantics, a verb denotes a property of individuals, whereas in event-based semantics, it denotes a relation between individuals and events. So, in her theory of Lexical Cumlativity, a verb denotes a relation between plural individuals and plural events. Now, we must prepare a place for atomic events and plural events in our domain. Following a convention, I introduce type v for events, assume a domain of event D_v. As D_e, D_v is closed under sum forming operation +. As *-operators are generalized so that it can pluralize a relation (and potentially higher order relations), a relation between individuals and events is pluralized. It is illustrated as follows.

(13) a. [[verb]] = {<John, e_1>, <Mary, e_2>, …}
     b. [[*verb]] = {<John, e_1>, <Mary, e_2>, <John+Mary, e_1+e_2>, …}

She assumes that nouns, verbs and thematic predicates such as *agent* and *goal* are also plural from the very start. As a result, we can obtain a cumulative reading without any effort.

(14) a. Two children lifted two boxes.
     b. $\exists e. \exists x. \exists y. [*child(x) \ \& \ |x| = 2 \ \& \ *agent\ (x)(e) \ \& \ *box\ (y) \ \& \ |y| = 2 \ \& \ *lift\ (y)(e)]$

Note that marking * on lexically plural predicates is just a notational matter and does not imply an application of a pluralizing operation. Again, they are plural from the very start. Importantly, the analysis does not distinguish one-time reading and repetitive reading. Also, the analysis does not distinguish collective readings and cumulative readings. Then, how about distributive readings? (14a) has two distributive readings, each of which is represented by (15a) and (15b).

(15) a. (2 children) *[lifted 2 boxes.]
     b. (2 boxes) *λ_1[2 children lifted t_1.]

They are described by the following denotation.

(16) a. *$\lambda x.\ \lambda e.\ \exists y.$ [*box (y) & $|y| = 2$ & *lift $(y)(e)$ & *agent $(x)(e)$]
  b. *$\lambda x.\ \lambda e.\ \exists y.$ [*child (y) & $|y| = 2$ & *lift $(x)(e)$ & *agent $(y)(e)$]

Thus, the analysis achieves a description of distributive readings by staring a verbal (phrasal) predicate. This is on a par with Landman's (1996, 2000) notion of distributivity as cumulativity with scope. However, pluralization of phrasal predicates must be done by a *-operator because they are by no means lexical. If so, though generalized *-operators can pluralize both lexical predicates and phrasal predicates, why do we need lexical plurality? On this point, Kratzer (2013) argues that the notion of generalized *-operators leads to a problem of overgeneration. First, she points out that singular indefinite objects fail to be within the scope of a *-operator.

(17) a. This intern guards a parking lot.
  b. She bounced a ball for 20 minutes.

(17a) describe a situation in which the intern habitually guards a single parking lot and (17b) describe a situation in which she bounced a single ball again and again. If *-operators are freely inserted, (17b) can produce (18b) from (18a). (18b) describes possibly repeated events in which more than a single ball is bounced.

(18) a. $\lambda e.\ \exists x.$ [ball (x) & *bounce $(x)(e)$]
  b. *$\lambda e.\ \exists x.$ [ball (x) & *bounce $(x)(e)$]

Thus, generalized *-operators are incompatible with the absence of the reading (18b). However, a *-operator is somehow needed for plurality of phrasal predicates. Kratzer's (2013) proposal is that the distribution of *-operators is closely tied with a presence of plural DPs: at the level where semantic interpretation takes place, the sister constituent of a plural DP is pluralized. A consequence of her account is that it should not be possible to simultaneously cumulate two non-event arguments. In her system, external arguments and applicative arguments are introduced by independent heads. Thus, pluralizing a predicate with more than one non-event arguments is not an option. To see it, consider the following example.

(19) These 5 teachers gave a bad mark to those 20 protesting students.

This sentence allows a cumulative interpretation. For instance, (19) is true in a situation in which each of the students got a bad mark from only one of the teachers. This example was, originally, a piece of evidence for generalized *-operators. Beck and Sauerland (2000) argues that the intended reading in (19) is derived only when we pluralize the two-place relation $\lambda x \lambda y \exists z$ [bad-mark (z) & gave-to $(y)(z)(x)$]. Thus, the reading requires the notion of a polyadic *-operator. However, Kratzer (2013) argues that the notion is unnecessary in the proposed event-based analysis. She alternatively proposes (20)

(20) *$\lambda y\ \lambda e\ \exists z$ [bad-mark (z) & *gave $(z)(e)$ & *goal $(y)(e)$]

The predicates 'gave' and 'goal' is plural via Lexical Cumulativity. Plurality of the predicate $\lambda y \lambda e$ is an instance of phrasal cumulativity and it is realized by moving the indirect argument above the direct object. As the indirect object is a plural DP, the sister constituent of it is pluralized. As a result, we obtain *$\lambda y \lambda e$. Thus, the relevant reading is derived without resorting

the polyadic *-operator. Next, I will see how the sister constituent of a plural DP is pluralized.

2.4.   Classifiers and Plurals

In this subsection, I see Kratzer's (2013) DP-internal structure of English, in which the sister constituent of a plural DP is pluralized. Primarily, there are two characteristics in her DP-internal structure. First, the [singular] feature is eliminated from her theory. Second, there are two positions for [plural] features in a DP. One is above nouns and the other is above Ds. Ds' [plural] feature functions as a *-operator for phrasal predicates. I examine them one by one.

First, she takes a position that there is no number feature [singular]. Though she emphasizes how significant the cumulativity universal is, singular count nouns are obvious counterexamples to this universal. To maintain the universal, she points out that count nouns are not simple but complex when we see them. Her point is that count nouns consist of a root and a piece of nominal inflection. Based on this assumption, there are two possible courses to take. First one is to argue that the denotation of common noun roots is predicative and *transnumeral* (Müller 2000, Rullmann and You 2006); they are number-neutral predicates. In this way, the cumulativity universal is satisfied. The other way is to argue that the denotation of common noun roots is referential and refers to kind (Krifka 1995, Yang 2001). As common noun roots are not predicative, the universal is maintained. These roots are, then, turned into predicates by nominal inflection. Kratzer's (2013) choice is the second way because this account does not specify how the denotation of noun roots is individuated or portioned. So, she suggests that English has a multiply ambiguous non-overt classifier, following Krifka (1995). For example, consider a noun root √*zebra*. It denotes a kind of zebra. Further, it is a role of non-overt classifiers to turn it into a predicate as in (21).

(21) a. $[[CL_{ind}]] = \lambda x \lambda y$ [kind (x) & individual (y) & $y \sqsubseteq x$]
     b. $[[CL_{kind}]] = \lambda x \lambda y$ [kind (x) & kind (y) & $y \sqsubseteq x$]

According to (21a), the word "zebra" is a singular predicate which is made from √*zebra* and an incorporated classifier. This is the reason of ambiguity between an individual and a subkind.

(22) a. This zebra has not been fed.
     b. This zebra is almost extinct.

What is pluralized, then, is this singular predicate made through the incorporation of a non-overt classifier into a noun root. If it is a job for the classifiers to provide a singular predicate, [singular] feature has no semantic job to do. She assumes the structure in (23) for English DPs.

(23)

([plural])

classifier   √*zebra*   (Kratzer 2013)
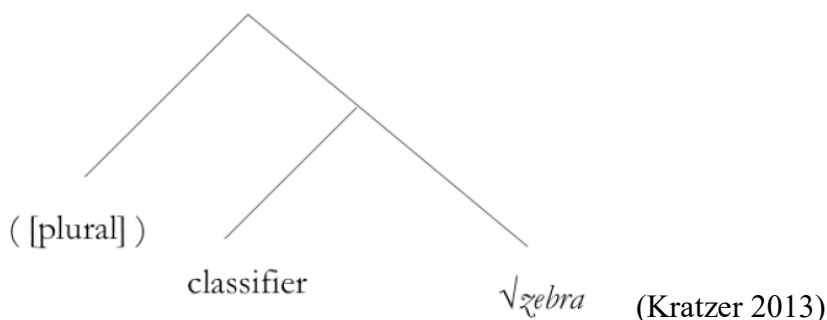
In this way, Kratzer (2013) eliminates the [singular] feature from her theory.

Next, she argues that there are two positions for the [plural] feature in a DP. Following Sauerland (2005), a plural DP is built from a determiner and a plural noun, each of which has its own number projection. Thus, (23) is revised as (24).

(24)

```
              /\
             /  \
    [plural]/    \
           /\     \
          /  \     \
         D    \     \
              /\     \
             /  \     \
    [plural]/    \     \
           /      \     \
          /       /\     \
         /       /  \     \
        classifier   N
```

(Kratzer 2013)
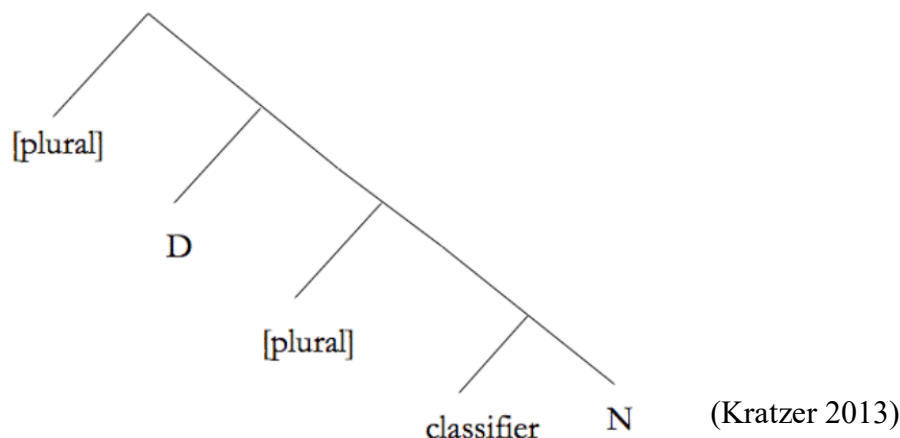
The higher [plural] is there even when a plural DP does not contain any plural NP. Consider a plural DP "Spencer and Webster." This DP triggers verbal agreement and have a distributive reading exactly like other plural DPs. Sauerland concludes that the higher [plural] is responsible for these properties. Then, Kratzer's (2013) point is that the higher [plural] feature is the source of phrasal plurality. Her story is as follows. The higher [plural] is not interpretable within DP, whereas the lower [plural] pluralizes a noun. As the higher [plural] is not interpretable within DP, it must move out of its DP before it is assigned semantic interpretation. Moving as little as possible, the [plural] feature becomes a verbal inflectional head. This [plural] feature is interpreted as a *-operator. This is how phrasal plurality is obtained when there is a plural DP.

The story seems to account for the reason why phrasal plurality is obtained only when a given predicate is the sister of a plural DP. However, some points in her account must be clarified. First, the reason why the higher [plural] feature is not interpretable within a DP needs to be explained. Kratzer (2013) assumes that [plural] features carry a cross-categorial *-operator. So, there might be no semantic reason for the higher [plural] to move out of its DP. It might be the case that the higher [plural] feature have nothing to pluralize because the lower [plural] feature has already pluralized a noun. However, there is no semantic problem with pluralizing an already pluralized object, though the result is vacuous. As a result, the status of the higher [plural] feature remains quite ad hoc. Moreover, in this account, the higher [plural] feature is lowered as in (25).

(25)

```
            /\
           /  \
          /    \
        DP      \
                /\
               /  \
              /    \
       [plural]     \
          = *
```

Pluralized sister predicate

(Kratzer 2013)
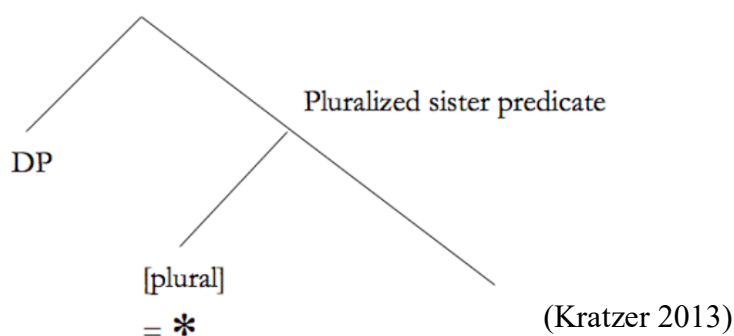
This kind of movement need to be justified carefully because the lowering operation is rarely observed in syntax. Thus, we had better find another system to guarantee that the sister predicate of a plural DP is pluralized. In the next section, I propose a modification of Kratzer's (2013) proposal so that it can overcome these problems.

3.   Cyclicity in semantics

In this section, I propose that the reason why the sister of plural DPs is pluralized is that D in English is a phase head. In this way, the problems above are answered without diminishing the coverage of Kratzer's (2013) theory of pluralities and events. In subsection 3.1, I review how the cycle is implemented in the notion of phase and in 3.2, I discuss the cyclicity in semantics and adopt a notion of phase. In subsection 3.3, I claim that a cyclic model of syntax-semantics interface dissolves all the theoretical problem found in Kratzer's (2013) theory of distributivity. In subsection 3.4., I discuss the mapping between syntax and semantics in the proposed account. Subsection 3.5. concludes this section.

3.1.   Phase as a cycle

In this subsection, I will briefly illustrate the notion of phase. Since the early days of generative grammar, it has been noticed that *discrete infinity* is one of the most important properties of natural language. As language expressions have potentially infinite length, the generative procedure that derives those expressions must allow *recursion* in the technical sense. Roughly speaking, for a generative procedure to derive an infinite expression by finite resources, the rules must be defined for the output of the same rule. In the early days, recursion is guaranteed by generalized transformation rules. However, generalized transformation rules are replaced by the phrase structure rules. It was made possible by the introduction of *cycle* into syntax. An early version of the implementation of cyclicity in syntax is as follows.

(26) <u>The Strict Cycle Condition (SCC)</u>: No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node. (Chomsky 1973)

This condition, for example, prevents (27d) from being derived.

(27) a. [$_A$ the election officials know [$_B$ to place who where on the ballot]]
b. [$_A$ ***who*** do the election officials know [$_B$ to place _ where on the ballot]]
c. [$_A$ who do the election officials know [$_B$ ***where*** to place _ _ on the ballot]]
d. ?? Who do the election officials know where to place on the ballot?

A wh-expression "who" is moved to the front of the cyclic domain labeled A. After that, another wh-expression "where" is moved to the front of the cyclic domain labeled B. This is a violation of SCC and thus (27d) is degraded. Note that SCC does not prevent long movement, but prevents movement in a lower domain after a higher domain has been accessed. At this stage, cycle and locality were indirectly related. In the Minimalist Program (Chomsky 1995 et seq), recursive structure is made by an operation *Merge*. This operation takes *n* objects and returns

the set of them. As it is generally considered that $n = 2$, we can define Merge as follows.

(28) Merge $(\alpha, \beta) = \{\alpha, \beta\}$

Allowing Merge to further apply to the previous outputs of Merge, Merge can generate infinite array of structured expressions. In this model, *phase* is the incarnation of cycle. The Phase Impenetrability Condition prohibits the application of an operation to the lower domain.

(29) <u>Phase Impenetrability Condition (PIC)</u>: In phase α with head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.

(Chomsky 2000)

Though the nature of the edge is still a matter of debate, phase edges function as an escape hatch for movement. So, every constituent moves to the edge if it moves out of the phasal phrase. So, the notion of phase plays the traditional role of cycle in the merge-based model of syntax. However, phase brings about a new aspect of cycle. One of them is the cyclicity of the interfaces. It has long been assumed that the phonological features are sent to the interface all at once (*single Spell-Out*). The cyclicity of the interface is to say that the phonological features are sent to the interface every time a phase is activated (*cyclic Spell-Out*). It leads to the possibility of unifying the phonological cycle and syntactic cycles. Therefore, it seems natural to ask whether the same effect is confirmed in the semantic interface. In the next subsection, I discuss the cyclicity in the syntax-semantics interface.

3.2. Phase in the syntax-semantics interface

In this subsection, I propose that a) the syntactic cycle affects semantic composition and, in addition, that b) this unification improves Kratzer's (2013) mechanism of phrasal pluralization. First, I consider that a domain qualifies as a phase if and only if the domain satisfies the PIC. In the mainstream theories of phases, phases are marked by some designated phase heads, e.g., C and *v*, and the phase head and its edge are accessible to further syntactic operation. However, in this paper, I adopt Bošković's (2014, 2016) proposal that the highest phrase in the extended projection of all lexical categories functions as a phase. So, there is always a nominal phase in his system and his idea is supported by cross-linguistic analysis on extraction and ellipsis. Also, Narita (2014) argues that transfer is triggered every time a phrase XP merges with another phrase YP. In his system, transferring the complement of DP, for example, is an available option. Importantly, his system allows the complement of KP in Japanese to be transferred and indeed, he support his claim by showing that the complement of KP disallows extraction from it in certain environment. Though their technical details vary, it is fair to say that the notion of nominal phase got empirical support to some extent. So, in a plural DP in English, the edge is D's [plural] feature as it is the highest element in the extended projection of the nominal domain. In the same vein, *v* the edge of *v*P as it is the highest element in the extended projection of the verbal domain if there is no higher verbal element above *v*.

3.3. Semantic cyclicity and Kratzer's (2013) system of phrasal pluralization

In this subsection, I argue that the problems found in Kratzer's (2013) mechanism of phrasal

pluralization are resolved by incorporating the notion of phase. In doing so, I also sketch out her system. To begin with, she argues that there are two positions for [plural] feature in a DP as follows.

(30)

[plural]

D

[plural]

classifier   N   (Kratzer 2013)
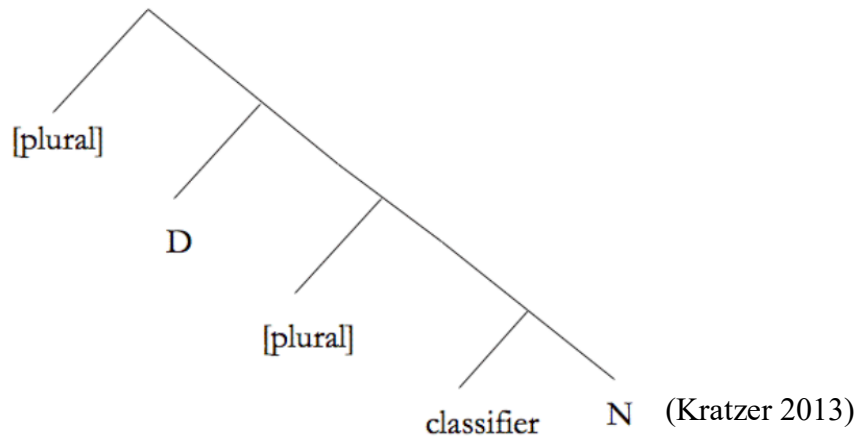
The higher [plural] is there even when a plural DP does not contain any plural NP. The higher [plural] is responsible for verbal agreement and distributivity. Second, the higher [plural] is not interpretable within DP, whereas the lower [plural] pluralizes a noun. As the higher [plural] is not interpretable within DP, it must move out of its DP before it is assigned semantic interpretation. Moving as little as possible, the [plural] feature becomes a verbal inflectional head. This [plural] feature is interpreted as a *-operator. However, it is problematic that the higher [plural] feature is always uninterpretable within a DP and the higher [plural] feature is lowered. These problems, however, disappear when we adopt semantic cyclicity. Specifically, [plural] feature of D is always at the phase edge and thus interpreted external to the DP. As [plural] feature in D is the highest element in the extended projection of the nominal domain, the complement of D's [plural] feature is transferred as follows.

(31)   DP

[plural]   Transfer domain

D

[plural]

classifier   N
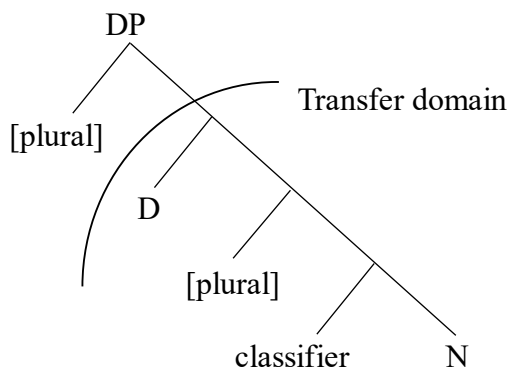
As a result, the low [plural] of N is always within the transfer domain and the high [plural] of D is always at the phase edge. Thus, the first problem is resolved. The high [plural] is interpretable within its DP in principle, but its position prevents it from being interpreted within the DP. On this point, I assume that an XP is reduced to X when its complement is transferred (Epstein 2009, Narita 2014, Lohndal 2014). Thus, the DP in (31) is reduced to [plural] feature. This reduced [plural] feature merges with *v*P as in (32). I abstract away from the issue of the internal argument in (32) to focus on the external argument.

(32)

[plural]   *v*P   Transfer domain

v   VP

V   IA
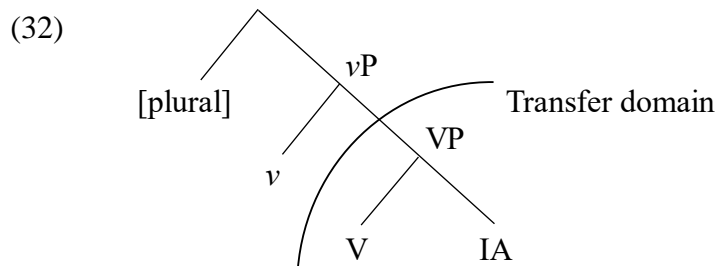
Then, as is desired, [plural] feature becomes the sister of *v*P and functions as a *-operator. Also, as *v* is the highest element in the extended projection of the verbal domain, the complement of *v* is transferred as in (32). So, the theoretical problems in Kratzer (2013) are resolved in the cyclic model of syntax-semantics interface. In the next subsection, I will propose a mapping rule which is necessary in this cyclic syntax-semantics interface.

3.4.   Mapping system

In this subsection, I discuss how transferred objects are mapped to semantic objects. As each syntactic structure is assigned its interpretation in cyclic fashion, their relation should be secured. A precursor of a cyclic mapping from syntax to semantics is found in Lohndal (2014). In his system, each transferred object is mapped to a conjunct in a Neo-Davidsonian logical form and conjoined in the semantic component. To achieve this, he assumes the rule of *Thematic Integration* as follows.

(33) **Thematic Integration**

H   DP   →Spell-Out→   R(e, DP).   (Lohndal 2014)

He assumes that H is one of Voice, F, Appl and so on. R is one of Agent, Theme, Experiencer and so on. 'F' is a newly proposed functional category that introduces the Theme. In this system, for example, the sentence in (34a) has a tree structure given in (34b), and has the logical form in (34c). [1]

(34) a. The student read the books.
     b.

VoiceP

The student   Voice   FP   Transfer domain

the books   F   VP   Transfer domain

read   (Lohndal 2014)

---
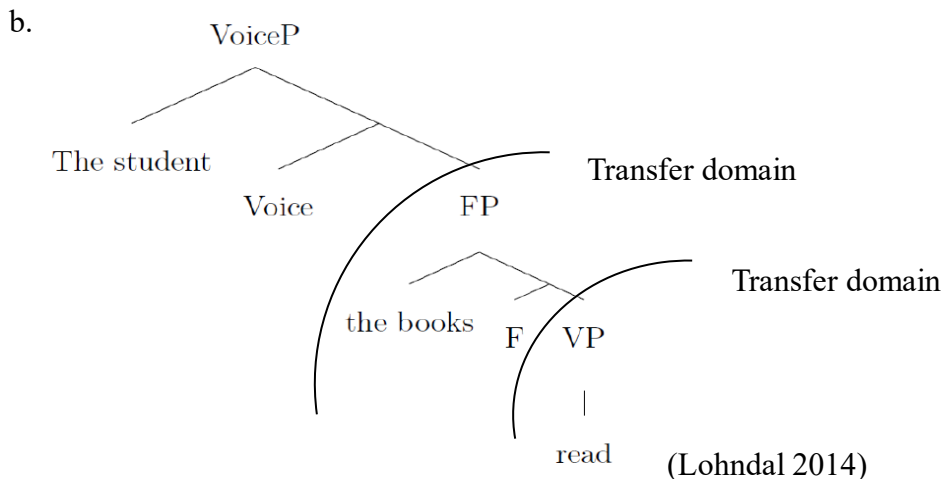
[1] In Lohndal's (2014) system, transfer is triggered every time {XP, YP} schema met.

    c. ∃*e*[kick(*e*)] & ∃*e* [Agent (*e*, the student)] & ∃*e* [Theme (*e*, the books)]

In these examples, each transferred structure is mapped to a thematic relation and its event variable is bound via *Existential closure*. However, it must be guaranteed that these event variables are synonymous to each other. Here, I propose an alternative. I adopt a slightly modified notion of *choice function* (Reinhart 1997, Winter 1997 among others). A choice function *f* is defined in (35):

(35) The choice condition: A function *f* is a choice function (i.e. CH (*f*) holds) *only if* for every nonempty predicate *P*, CH (*P*) is defined and it is in the extension of *P* (i.e. *P*(*f*(*P*)) holds).
(Winter 1997)

The notion of choice function is first proposed to account for the exceptional wide scope readings of indefinites. I generalize this notion of choice function so that event predicate is also included in the range of choice function *f*. My proposal is that choice function mechanism applies every phase level. For example, see the derivation of a distributive reading of (36).

(36) The students read a book.

The object plural DP 'a book' is mapped as follows.[2] As the article 'a' is the highest element within the nominal domain, the D complement is transferred.

(37) a.

```
                    DP
                   /  \
                  /    \____ Transfer domain
                 /      \
          D     /        \
          a    /          \
              /            \
          [plural]          \
                 \           \
                  \           \
               classifier      N
                              book
```
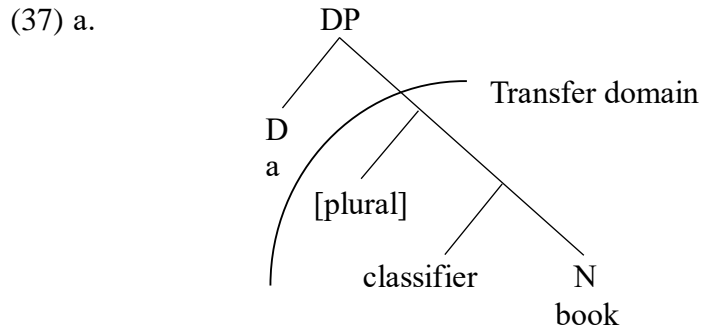
    b. Transfer domain: λ*x* [book (*x*)]

When the transfer domain of (37) is sent to the semantic component, the choice function mechanism applies and (38) derives.

(38) *f* [CH (*f*) & *f*(book)]

Following Reinhart (1997) and Winter (1997), I assume existential closure of the free variable *f* is applied every compositional level. However, how are the transferred subject and V combined? On this point, I stipulate that *f*(*P*) such that *P* is within a preceding phase saturate a predicate within the next phase.[3] Thus, in this case, *y* is within the λ*y*. λ*e* [*read (*e*, *y*)] is saturated by *f*(book). To express this, I use the following notation:

---

[2] I take a position that the indefinite article 'a' is semantically vacuous.
[3] For a precise formulation of this process, I will adopt the plural pronoun/definite description [*Ê*: <Φ>ⁿ] (Schein 1993, 2012) and *accessibility constraint* in *discourse representation theory* (Kamp and Ryle 1993), though I do not go into detail here.

(39) λe [*read (e, y)] & [y:f(book)]

Next, the subject plural DP 'the students' is derived as follows.

(40) a.
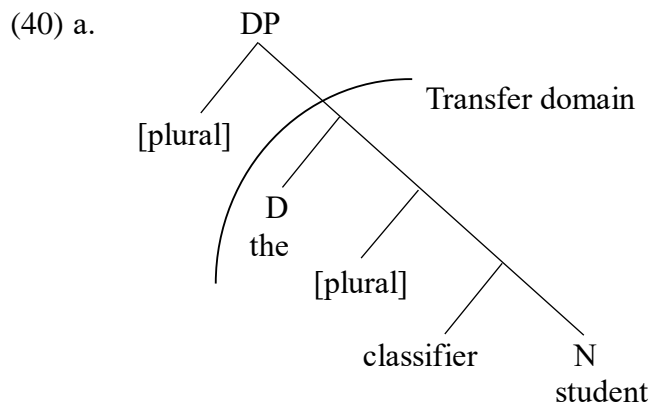
b. Transfer domain: λx [*student (x) & $|$*student$| = 1$][4,5]

After the transfer in (40), [plural] feature is introduced to a clausal spine. As v is the highest element in the extended projection of the verbal domain, the complement of v is transferred as follows.

(41)
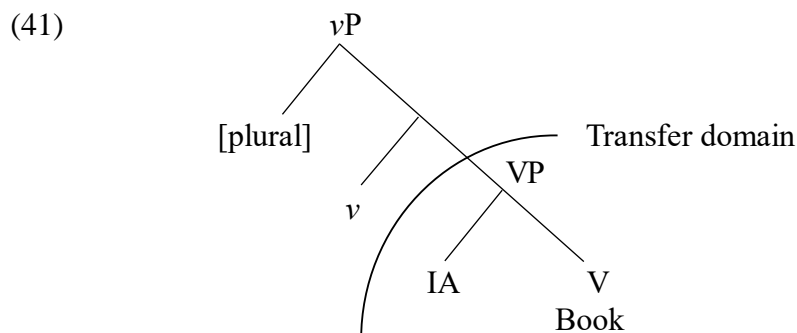
When C merges subsequently, its complement is transferred. The resulting formula is as in (42).

(42) a. *λx. [*agent (e, x) & $\exists f_2$ [e: *read (e, y) & [y:f_2(book)]]] $(x$:f_1(*student))
    b. $\exists f_1$ *[*agent (e, x) & [x:f_1(*student)] & $\exists f_2$ [e: *read (e, y) & [y:f_2(book)]]][6]

Note that if existential closure applies to the choice function variable f_2 after the C complement is transferred, a wide scope reading of 'a book' is obtained. Coupled with this mapping system, Kratzer's (2013) system is modified so that it allows cyclic mapping. In the next section, I argue that this cyclic model has better explanation for the event-related readings in English and Japanese.

---

[4] I adopt a version of Russellian analyses, here: 'the' denote a predicate modifier of $<<e, t>, <e, t>>$ type, which imposes a uniqueness presupposition by requiring the set to be a singleton.

[5] As is mentioned in 2.3., the extension of a cumulative predicate is a set comprising of atomic individuals and plural individuals. So, f(*P) can potentially refer to a plural individual. Also, it's possible that a set is a singleton set comprising one and only one plural individual.

[6] Though *[Ψ] may appear unfamiliar to many readers, I use this notation to indicate that this formula is derived from *λx [$P(x)$] in order to signify phrasal plurality.

4.   *Applications*

In this section, I show some cases in which the proposed account offers a better account than Kratzer (2013) does. Subsection 4.1. treats event-related readings of numeral quantifiers in English. In subsection 4.2., I show floating numeral quantifiers in Japanese allow only event-related reading. In the subsection 4.3., I relativize the proposed account so that it also accounts for Japanese data.

4.1.   *Event-related readings of English numerals*

Numeral quantifiers in English can measure the number of events as follows.

(43) Four thousand ships passed through the lock last year. (Krifka 1990)

Example (43) can be uttered in a situation where there were 4000 passing through events of just a single ship last year. In this reading, the DP internal numeral 'four thousand' is interpreted externally to the DP. I call this reading an *event-related reading*, which is contrasted to an *object-related reading*, in which the number of object e.g., ships, is counted. In the proposed account, this ambiguity is accounted for with just one assumption. That is, there can be two distinct positions for numerals in Krazter's DP structure: N's [plural] feature and D's [plural] feature. Then, in the proposed account, N's [plural] feature is always noun-internal and D's [plural] feature is always noun-external. So, this object-related/event-related distinction is accounted for as a direct reflection of their position. Specifically, the subject DP is as in (44) when it induces an object-related reading and as in (46) when it induces an event-related reading.

(44) Nominal domain: Object-related reading

a.
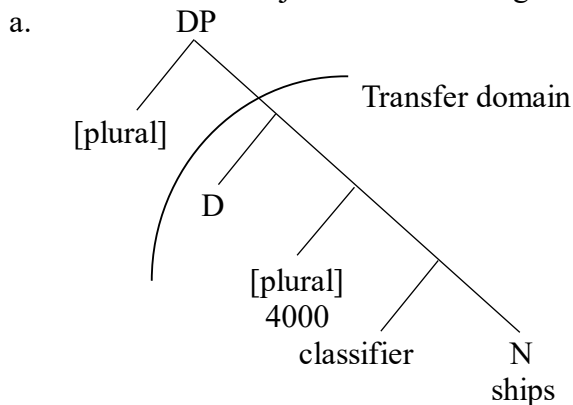
b. Transfer domain: f [CH (f) & f(*ship) & |f(*ship)| = 4000]

(45) Verbal domain: Object-related reading

a.
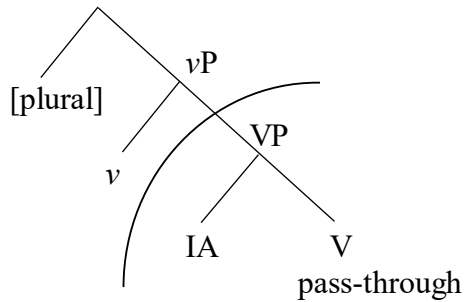
b. VP: [e: *pass-through (e, y) & [y: f(lock) & |lock| = 1]]

c. {DP, vP}: $\exists f_1$ *[*agent (e, x) & [x: f_1(*ship) & |f(*ship)| = 4000] &
[e: *pass-through (e, y) & $\exists f_2$ [y: f_2(lock) & |lock| = 1]]]

(46) Nominal domain: Event-related reading

a.
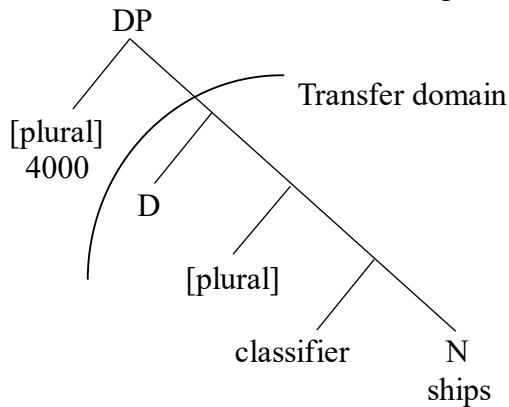
b. Transferred domain: f [CH (f) & f(*ship)]

(47) Verbal domain: Event-related reading

a.
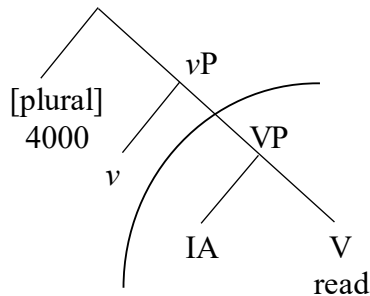
b. [e: *pass-through (e, y) & [y: f(lock) & |lock| = 1]]

c. {[plural], vP}: $\exists f_1$ *[*agent (e, x) & [x: f_1(*ship)] & [e: *pass-through (e, y) &
|e| = 4000 & $\exists f_2$ [y: f_2(lock) & |lock| = 1]]

This kind of treatment is not allowed in Kratzer's (2013) theory of non-cyclic mapping. However, in this account, it is allowed as a natural consequence.

4.2.   *Floating numeral quantifiers in Japanese*

In this section, I discuss floating numeral quantifiers[7] (FNQ) in Japanese and argue that their interpretations are rigidly event-related. First, unlike English and other European languages, Japanese allows numeral quantifiers to float as follows.

(48) a. San-nin-no   gakusei-ga   hasi-ta. (non-floating NQ)[8, 9]
         3-CL$_{human}$-no student-Nom run-Past[10]
     b. Gakusei-ga   san-nin   hasi-ta. (floating NQ)
         student-Nom 3-CL$_{human}$ run-Past
         'Three students ran.'

Nakanishi (2007) argues that FNQs quantify event. To confirm this, she uses a combination of the predicate 'koros-u' (kill) and a proper name. Since it is impossible to kill a person repeatedly, the combination is incompatible with plural events as in (49).

(49) #Gakusei-ga   kinoo   <u>san-kai</u>   Peter-o   korosi-ta.
        student-Nom yesterday <u>3-CL$_{event}$</u> Peter-Acc kill-Past
        'Students killed Peter three times yesterday.'

(Nakanishi 2007)

Then, compare (50a) and (50b).

(50) a. San-nin-no   gakusei-ga   kinoo   Peter-o   korosi-ta. (non-floating NQ)
         3-CL$_{human}$-no student-Nom yesterday Peter-Acc kill-Past
     b. #Gakusei-ga   kinoo   san-nin   Peter-o   korosi-ta. (floating NQ)
         student-Nom yesterday 3-CL$_{human}$ Peter-Acc kill-Past
         'Students killed Peter three times yesterday.'

(Nakanishi 2007)

The unacceptability of (50b) indicates that 'san-nin' (3-CL$_{human}$) in (50b) measures events. As (50a) allows both an object-related reading and an event-related reading, (50a) sounds grammatical under the scenario in which three students together kill Peter. However, (50b) cannot express this reading, as is indicated by the unacceptability of (50b). Also, there are cases in which multiple NQs co-occur with a host nominal as in (51).

---

[7] I refer to the unit of a numeral e.g., 'san' (three), and a classifier e.g., '-nin' (a classifier for human-beings), as a *numeral quantifier* (NQ) and refer to a nominal predicate that is quantified over by an NQ e.g., 'gakusei' (student), as a *host nominal*.
[8] Following Watanabe (2006), I assume 'no' that follows non-floating NQs is a linker which is inserted in the morphological component. So, I just write 'no' in gloss.
[9] Actually, there is prenominal and postnominal distinction among non-floating NQs, though I do not discuss this issue in this paper. For their distinction, see Nakamura (to appear).
[10] I use subscript $_{CL}$ to express that a given item is a classifier.

(51) [Naran-de   hashit-tei-ta   su-dai-no   torakku]-ga   gaadoreeru-ni
    side-by-side   run-Prog-Past   some-CL$_{car}$-no   truck-Nom   guardrail-at
   san, yon-dai   butsukat-ta.
    3, 4-CL$_{car}$ bump-Past
   'Three to four of the trucks that ran side by side bumped the guardrail.' (Inoue 1978)

Inoue's (1978) point is that 'san, yon-dai' (3, 4-CL$_{car}$) in (51) should be analyzed as an adverb because it has no base-position at the NP. If so, it predicts that this kind of FNQs quantify events. This prediction is borne out as follows.

(52) #[watasi-ga   kinoo   at-ta   suu-nin-no   gakusei]-ga
      I-Nom yesterday meet-Past some-CL$_{human}$-Gen   student-Nom
   san-nin   Peter-o   korosi-ta.
   3-CL$_{human}$ Peter-Acc   kill-Past
   'Three of the students who I met yesterday killed Peter.'

Example (52) forces the interpretation where there are three events of killing Peter. However, such a situation is impossible in the real world. If (52) allowed an interpretation that a group of three student killed Peter in cooperation, it would sound perfect because there is a single event of killing. The fact that (52) sound strange means that this option is unavailable. Thus, the FNQ 'san-nin' (3-CL$_{human}$) in (52) measures events. Note that there is a partitive relation between 'suu-nin' (some-CL$_{human}$) and 'san-nin' (3-CL$_{human}$). There are further cases which involves *Event Partitioning Quantifiers* (EPQs) (Nakamura 2016, 2017). EPQs are defined as a complex NQ consisting of an NQ 'hito-ri' (1-CL$_{human}$) followed by an arbitrary NQ. In this sense, there are two host nominals per EPQ because an EPQ contains two NQs. EPQs are exemplified below.

(53) a. Kodomo-ga   okasi-o   hito-ri-huta-tu   ka-ta.
     Child-Nom candy-Acc 1-CL$_{human}$-1-CL$_{thing}$ buy-Past
    'Every child bought two candies.'
   b. Gakusei-ga   hito-ri-san-kai   hasi-ta.
    Student-Nom 1-CL$_{human}$-3-CL$_{event}$ run-Past
    'Every student ran three times.'

As is reported in Nakamura (2016, 2017), an EPQ can co-occur with other NQs within nominal domain as follows.

(54) San-nin-no   gakusei-ga go-satu-no hon-o   hito-ri-ni-satu   kat-ta.
   3-CL$_{human}$-no student-Nom 5-CL$_{book}$-no book-Acc 1-CL$_{human}$-2-CL$_{book}$ buy-Past
   'Each of three students bought two of the five books.'

The most important characteristic of EPQs is that the object functions as a range of the theme of the events. For example, in (54), there are three students and they chose two books each from the five books. In other words, the NQ 'go-satsu' (5-CL$_{book}$) specifies the type of books and the NQ 'ni-satsu' (2-CL$_{book}$) specifies each tokens of the books for each event. The same thing is observed for the subject when an EPQ appears with an NQ whose classifier is '-kai.' (CL$_{event}$)

(55) San-nin-no gakusei-ga go-satu-no   hon-o   hito-ri-ni-satu san-kai kat-ta.
     3-CL$_{human}$-no student-Nom 5-CL$_{book}$-no book-Acc 1-CL$_{human}$-2-CL$_{book}$ 3-CL$_{event}$ buy-Past
     a. 'Each of three students bought two of the five books three times.'
     b. ?'There are three event of buying. Each agent of the events is chosen once from the
        three students and each theme of the events is chosen twice from the three books.'

The reading (55a) is run-of-the-mill repetitive reading. What is relevant for the discussion is
the reading (55b). Though it is a little bit difficult to take, this reading is guaranteed by native
speakers of Japanese. The existence of the reading (55b) indicates that the subject as well as the
object functions as a range of possible participants of each event. To guarantee this reading, it
is necessary to interpret the plural individual in the nominal domain and the plural individuals
in the verbal domain, independently. On this point, the notion of cycle plays a crucial role. Once
the phrase 'go-satu no hon' is interpreted, this interpretation must not be altered by further
application of semantic operations. The same thing applies to 'san-nin no gakusei' in the reading
(55b). Additional support for this kind of cyclicity is provided by the fact that the object is
always non-quantificational: it takes neither a narrow scope reading nor a wide scope reading
with respect to any quantificational item. Summing up, Japanese FNQs are rigidly event-related
and they allow co-occurrence of another non-floating NQ.

4.3.   Application to Japanese

In this subsection, I discuss how the proposed model is applied to Japanese and how it accounts
for the data described in the previous subsection. First, I assume Japanese lacks D (Fukui 1995,
Kuroda 1992). Instead, I assume functional category K is phase head in Japanese. K head is
spelled out as a case-particle (Takahashi 2011, Narita 2014, Bošković's 2014, 2016). So,
Japanese nominal domain has the following structure.

(56)   KP
                      ╱ ╲
                    ╱   ╲ Transfer domain
            [plural]   ╲
                    ╲   ╲
                      K   ╲
                        ╲   ╲
                    [plural]   ╲
                          ╲   ╲
                      classifier N
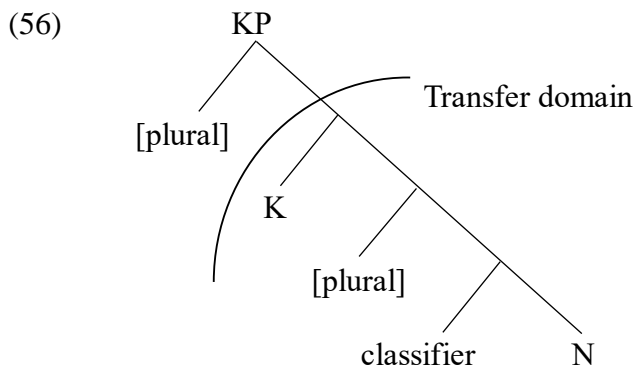
The presence of [plural] above K is shown in parallel to D's as follows.

(57) Taro-to-Hanako-ga   go-satu-no-hon-o   ka-ta.
     Taro-and-Hanako-Nom 5-CL$_{book}$-no-book-Acc buy-Past
     'Taro and Hanako bought two books.'

Example (57) allows a cumulative reading and a distributive reading, which suggests that the
phrasal predicate that is the sister of 'Taro-to-Hanako' (Taro and Hanako) is pluralized. Then,
the rigid event-relatedness is straightforwardly derived from the position of FNQs. Specifically,

non-floating NQs have their base position at the [plural] of N, whereas floating NQs directly merges to KP. As non-floating NQs have two distinct positions within the KP, they are ambiguous between an object-related reading and an event-related reading. On the contrary, floating NQs have unique position at the edge of KP phase. So, they allow only an event-related reading. This suggestion predicts that there can be cases in which both position is filled with an NQ as in (58).

(58)

[tree diagram: KP with branches to [plural] FNQ, K, and Transfer domain containing [plural] non-floating NQ, classifier, N]
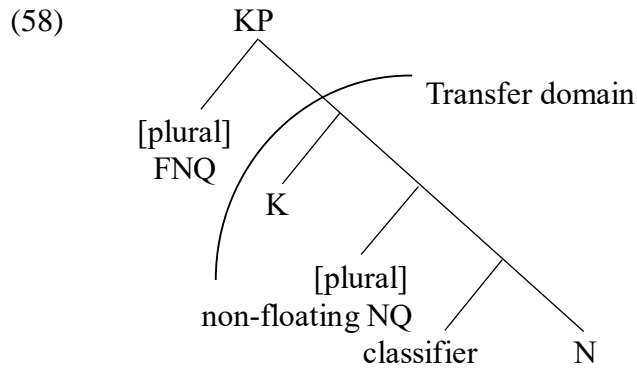
For example, the KP 'narande hari-te-ita suu-dai no torakku-ga san, yon-dai' is as follows.

(59) a.

[tree diagram: KP with branches to [plural] San, yondai, K ga, Transfer domain containing RC narande hari-te-ita, [plural] suu, classifier dai, N torakku]
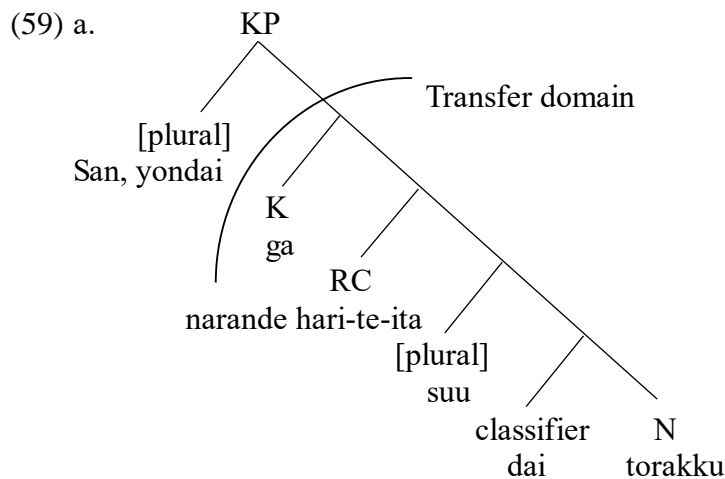
b. Transfer domain: *f* [*f*(*track) & *running-in-parallel (*f*(*track)) & |*f*(*track)| =some]

Then, after the transfer domain is given the denotation above, the FNQ 'san, yon-dai' merges with *v*P as follows. Note that agent relation has cumulative denotation.

(60) a.

[tree diagram: vP with branches to [plural] san, yon-dai, v, VP containing IA, V butsuk-]
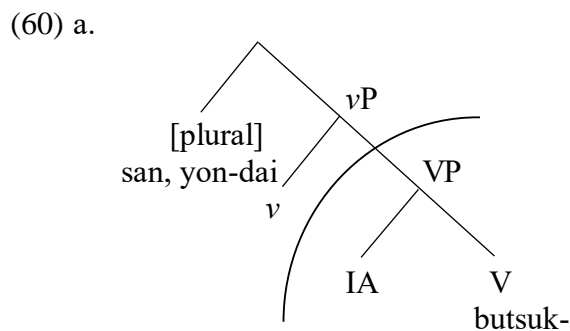
b. vP: *λx [*agent (e, x) & [e: *bump (e, y) & [y: f (*guardrail)]

c. {[plural], vP}: $\exists f_1$ *[*agent (e, x) & [x:f_1(*track) & *running-in-parallel $(f$(*track)) & $|f$(*track)$|$ =some] & [e: *bump (e, y) & $|e|$ = 3 or 4 & $\exists f_2$ [y:f_2(*guardrail)]]]

So, the proposed model makes a correct prediction with respect to the Japanese data. The scope-less interpretation of the non-floating NQ is forced because D's position of [plural] is already filled with the FNQ. As has been illustrated above, modified version of Kratzer's (2013) theory account for the rigidly event-related status of FNQs in Japanese.

5.   Conclusion

In this paper, I proposed a modification of Kratzer's (2013) theory with a cyclic mapping between syntax and semantics. In the proposed account, the highest element within an extended projection of a lexical category is the phase edge and the sister of the edge is transferred. So, the D's [plural] feature is always at the edge of the nominal phase and so interpreted as a *-operator within the verbal domain. In this approach, ad hoc status of the D's [plural] feature and its lowering are not necessary. Furthermore, the proposed account straightforwardly account for event-related readings of numerals in English and rigidly event-related nature of floating numeral quantifiers in Japanese. So, the proposed model covers wider empirical phenomena without postulating an ad hoc feature or a problematic syntactic operation.

Takanobu Nakamura
Sophia University
paradigm.shift.forward@gmail.com

References

Beck, S. & U. Sauerland (2000). Cumulation is needed: A reply to Winter. *Natural language semantics* 8:4, pp. 349-371.

Bošković, Ž. (2014) Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45:1, pp. 27-89.

Bošković, Ž. (2016) Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47:1, pp. 1-33.

Cable, S. (2014) Distributive numerals and distance distributivity in Tlingit (and beyond). *Language* 90:3, pp. 562-606.

Chomsky, N. (1973) Conditions on transformations. Anderson. S. R. & P. Kiparsky (eds.), *A Festschrift for Morris Halle*, Holt, New York, pp. 232–286.

Chomsky, N. (1995) *The Minimalist Program*. MIT Press, Cambridge: MA.

Chomsky, N. (2000) Minimalist Inquiries: The Framework. Martin. R., D. Michaels & J. Uriagereka (eds.), *Step by Step: Minimalist Essays in Honor of Howard Lasnik*, MIT Press, Cambridge: MA, pp. 89–155.

Epstein, S. D. (2009) The Unification of Theta Relations: How TRANSFER Renders Spec vp a Theta Marked Complement. [Ms]. University of Michigan.

Fukui, N. (1995) *Theory of projection in syntax*. CSLI Publications, Stanford: CA.

Inoue, K. (1978) *Nihongo-no bunpou kisoku* [*Grammatical rules of Japanese*.] Taishuukan, Tokyo.

Kratzer, A. (2013) On the plurality of verbs. Dölling. J., T. Heyde-Zybatow & M. Schäfer (eds.), *Event Structure in Linguistic Form and Interpretation*, De Gruyter, Berlin, pp. 269–300.

Krifka, M. (1990) Four Thousand Ships Passed Through the Lock: Object-Induced Measure Functions on Events. *Linguistics and Philosophy* 13, pp. 487–520.

Krifka, M. (1995) Common Nouns: A Contrastive Analysis of Chinese and English. Carlson. G. N. & F. J. Pelletier (eds.), *The Generic Book*, The University of Chicago Press, Chicago, pp. 398-411.

Krifka, M. (1998) The origins of telicity. Rothstein. S (eds.), *Events and grammar*, Springer, Netherlands, pp. 197-235.

Kuroda, S.-Y. (1992) *Japanese Syntax and Semantics*. Springer, Netherland.

Landman, F. (1996) Plurality. S. Lappin (eds.), *Handbook of Contemporary Semantics*, Blackwell, Oxford, pp. 425-457.

Landman, F. (2000) *Events and Plurality*. Kluwer, Dordrecht.

Link, G. (1983) The logical analysis of plurals and mass terms: a lattice- theoretic approach. Portner. P & B. H. Partee (eds.), *Formal Semantics: The Essential Readings*, Blackwell, Oxford, pp. 127-147.

Lohndal, T. (2014) *Phrase Structure and Argument Structure: A Case Study of the Syntax-semantics Interface*. Oxford University Press, Oxford.

Müller, A. (2000) The Expression of Genericity in Brazilian Portuguese. Kusumoto. K. & E. Villalta. (eds.), *UMOP 23: Issues in Semantics*, GLSA, Amherst: MA, pp. 137-154.

Nakamura, T. to appear. Predicative/Quantificational Numeral Quantifiers in Japanese. *Proceedings of WAFL 13*.

Nakamura, T. (2016) Events, distributivity and Numerals: Evidence from Event Partitioning Quantifiers in Japanese. Kim. T. S. & S. Ha. (eds.), *Proceedings of the 18th Seoul International Conference in Generative Grammar*, Sogang University, South Korea.

Nakamura, T. (2017) The Semantic Ontology of Agent and Theme: A Case Study with Event Partitioning Quantifiers in Japanese. Djärv. K. & A. G. Davies. (eds.), *U Penn Working Papers in Linguistics* 23:1, University of Pennsylvania, Philadelphia, pp. 171-179.

Nakanishi, K. (2007) *Formal properties of measurement constructions*. De Gruyter, Berlin.

Narita, H. (2014) *Endocentric structuring of projection-free syntax*. John Benjamins, Amsterdam.

Reinhart, T. (1997) Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and philosophy* 20, pp. 335-397.

Rullmann, H. & A. You. (2006) General Number and the Semantics and Pragmatics of Indefinite Bare Nouns in Mandarin Chinese. Von Heusinger. K. & K. Turner. (eds.), *Where semantics meets pragmatics*, Brill, Leiden, pp. 175-196.

Sauerland, U. (2005) On the Semantic Markedness of Φ -Features. [Ms] ZAS Berlin.

Sauerland, U., J. Anderssen & K. Yatsushiro. (2005) The plural is semantically unmarked. Kesper. S. & M. Reise. (eds.), *Linguistic Evidence*, Mouton de Gruyter, Berlin, pp. 409-430.

Schein, B. (1993) *Plurals and Events*. MIT press, Cambridge: MA.

Schein, B. (2012) Event Semantics. Russell. G. & D. G. Fara. *Routledge Companion to Philosophy of Language*, Routledge, Oxford, pp. 280-294.

Takahashi, M. (2011) *Some consequences of Case-marking in Japanese*. Doctoral dissertation, University of Connecticut, Storrs.

Watanabe, A. (2006) Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language & Linguistic Theory* 24:1, pp. 241-306.

Winter, Y. (1997) Choice functions and the scopal semantics of indefinites. *Linguistics and philosophy* 20, pp. 399-467.

Yang, R. (2001) *Common Nouns, Classifiers, and Quantification in Chinese*. Doctoral Dissertation, Rutgers University.

Conjunctive or Disjunctive? On the Syntax/Semantics of *-toka* and *-tari* in Japanese

Ryoichiro Kobayashi and Ryan Walter Smith

This paper investigates the nature of the Japanese non-exhaustive particles *-toka* and *-tari*. At first glance, the distribution of these particles is very similar to that of the focus particles *-mo* 'also' and *-sae* 'even': they are both used as coordinators and stand-alone particles, are in-compatible with topics (anti-topical), and induce intervention effects. However, while *-sae* and *-mo* carry presuppositions that project out of non-veridical contexts, *-toka* and *-tari* do not. Instead, they receive inclusive disjunctive interpretations in this context. We analyze *-toka* and *-tari* as items that introduce alternatives, which, once they expand into propositions, are either universally or existentially quantified depending on the veridicality or non-veridicality of their environment, respectively, and derive their coordinative use from their basic use as single particles.

1. Introduction

In Japanese, the particles *-toka* and *-tari* are used in veridical contexts to provide non-exhaustive conjunctions of nominal and predicative structures, respectively, as shown in (1).

(1) a. Taro -toka   Hanako -toka-ga   ki   -ta
     T.   -toka   H.   -toka-NOM   come -PAST
     'Taro, Hanako, and others came.'

   b. Ziro -ga   ringo -toka mikan   -toka-o   tabe -ta.
      Z.   -NOM apple -toka orange   -toka-ACC   eat   -PAST
      'Ziro ate an apple, an orange, etc.'

   c. Taro-wa   heya -o   soojisi -tari eigo   -o   benkyoosi -tari si -ta
      T.   -TOP room -ACC clean   -tari English -ACC study   -tari   do -PAST
      'Taro cleaned his room, studied English, and did other things.'

   d. Watasi -wa   Hanako -ga   kawaikat-tari, kasikokat -tari-site,   urayamasi -i.
      I   -TOP H.   -NOM cute   -tari smart   -tari-do   envious   -PRES
      'I'm envious of Hanako for being cute, smart, etc.'

In addition to their use as coordinators, *-toka* and *-tari* may appear as stand-alone particles attached to nominal and verbal elements respectively. In (2a), *Nihongo-toka* 'Japanese-toka' denotes a set of languages that John is likely to study, whereas *Eigo-o benkyoosi-tari* 'study-tari English' in (2b) denotes a set of predicates related to 'studying English'.

(2) a. John -wa   nihongo -toka-o   benkyoosi -ta.
       J.   -TOP   Japanese-toka-ACC   study   -PAST
       'John studied Japanese among other things.'

    b. Taro-wa   eigo   -o   benkyoosi -tari   su -ru.
       T. -TOP   English -ACC   study   -tari   do -PRES
       'Taro studies English among other things.'

We analyze *-toka* and *-tari* as items that introduce alternatives, which, once they expand into propositions, are either universally or existentially quantified depending on the veridicality or non-veridicality of their environment, respectively, and derive their coordinative use from their basic use as single particles. The rest of this paper is organized as follows. In section 2, we provide an in-depth description of the particles, hitherto less investigated in the literature. Comparison between *-toka/tari* and *-mo/sae* leads us to propose a unified syntactic structure, following Mitrović and Sauerland (2014) in section 3. Exploring its variable nature in veridical and non-veridical contexts, we further propose a semantic analysis of *-toka* and *-tari*, based on the semantics of alternatives. Section 4 concludes with some notes on the possible shortcomings of our analysis and some potential directions for addressing these shortcomings.

2. Non-exhaustive particles in Japanese
2.1 Similarities between -toka/tari and -mo/sae

At first glance, *-toka* and *-tari* seem to pattern very closely with the focus particles *-mo* and *-sae*. First, *-toka* and *-tari* can stand on their own, acting much like focus particles themselves, as illustrated in (3) and (4) below.

(3) a. John -wa   Nihongo-toka-o   benkyoosi -ta.
       J.   -TOP   Japanese-toka-ACC   study   -PAST
       'John studied Japanese among other things.'

    b. Taro-wa   Eigo   -o   benkyoosi -tari   su-ru.
       T. -TOP   English -ACC   study   -tari   do-PRES
       'Taro studies English among other things.'

(4) a. Hanako -wa   nigatena   burokkorii -sae /mo   tabe -ta.
       H.   -TOP   not.good.at broccoli   -even/also eat   -PAST
       'Hanako even/also ate broccolis, which she doesn't like very much.'

    b. Yuko -wa Hiroshimaben   -o   hanasi -sae /mo su-ru.
       Y.   -TOP Hiroshima.dialect-ACC speak   -even/also do-PRES
       'Yuko even/also speaks the Hiroshima dialect.'

Second, *-mo* and *-sae* can themselves act as coordinators in a manner similar to *-toka* and *-tari*, as demonstrated in (5) and (6). Although *-sae* does not seem to coordinate nominals, it is compatible with verbal coordination in (5b). Some might wonder whether *nusumi* 'rob' and *korosi* 'murder' in their *ren'yookei* 'conjunctive' form are actually nominalized verbs, since the *-i* inflectional ending tends to make derived nominals (Kageyama 1993). We argue that this is not the case, as *nusumi* and *korosi* in (5b) still take direct objects *kane-o* 'money-ACC' and *hiro-o* 'human-ACC' respectively.[1]

(5) a. Taro -mo Hanako -mo paatii -ni   ki   -ta
       T. -mo H.   -mo party -to   come -PAST
      'Taro and Hanako also came to the party.'

    b. Kare-wa   kane   -o   nusumi -mo/sae, hito   -o   korosi -mo/sae su -ru
       he -TOP money -ACC rob   -mo/sae   human -ACC murder -mo/sae do -PRES
      'He also/even robs money and murders people.'

(6) a. Ziro -ga   ringo -toka mikan   -toka-o   tabe -ta.   (=(1))
       Z.   -NOM   apple -toka orange -toka-ACC eat   -PAST
      'Ziro ate an apple, an orange, and other such things.'

    b. Taro-wa heya -o   soojisi -tari eigo   -o   benkyoosi -tari si-ta
       T.   -TOP room -ACC clean   -tari English -ACC study   -tari do-PAST
      'Taro cleaned his room, studied English, and did other things.'

Third, *-toka* and *-tari* are incompatible with the topical particle *-wa*, just as *-sae* and *–mo* are. All of these items are unacceptable when marked with *-wa* unless interpreted contrastively, as illustrated in (7). In this sense, they are all anti-topical items (Tomioka 2007). If *-toka* and *-tari* are anti-topical, then they would be expected to obtain illegitimate information structure, or cause intervention effects in *wh*-interrogatives (hitherto known as LF-intervention effects first pointed out by Hoji 1986).[2] This prediction is indeed borne out: the nominal particles induce intervention effects in (8) and (9).[3] Note that they become grammatical when the *wh*-element is overtly scrambled over the intervener, which is one of the well-known characteristics of intervention effects, as in (8b) and (9b).

(7) a. *Taro   -mo/sae -wa   ki   -ta.
        T.   -mo/sae -TOP   come -PAST
       'As for also/even Taro, he came.'

---

[1] The anonymous reviewer kindly mentioned that in many languages, derived nominals license accusative case on objects (Alexiadou et al. 2011). If this holds true in Japanese, the fact the nominals of the *ren'yookei* form in (5b) take direct objects would not be a conclusive argument for *-sae* and *-mo* to coordinate verbal phrases. As it is not uncontroversial, we leave open the discussion for future research. We would like to thank the reviewer for bringing up Alexiadou et al. (2011) to our attention.

[2] It is not uncontroversial whether the relevant intervention effects in Japanese are syntactic or pragmatic in nature. See Tomioka (2007) and Kobayashi (2017) for discussion and the references cited therein.

[3] Since it is difficult to use verbal particles in the environments where *wh* and foci induce intervention effects, we focus on nominal particles here.

b. *Taro -toka-wa   ki   -ta.   (^{ok}contrastive/*topic)
   T.   -toka-TOP   come -PAST
   'As for Taro etc., they came.'

c. *Soojisi-tari-wa   Taro -ga   si -ta.   (^{ok}contrastive/*topic)
   clean -tari-TOP   T   -.NOM   do -PAST
   'As for also cleaning, Taro did.'

(8) a. *?Hanako-mo/sae dare-o   home-ta   no?
       H.-mo/sae   who-ACC praise-PAST Q
       'Who did also/even Hanako praise?'

   b. Dare$_i$-o   Hanako-mo/sae t_i home-ta   no?
      who-ACC H.   -mo/sae   praise-PAST   Q
      'Who did also/even Hanako praise?'

(9) a. *?Taro -toka-ga   nani -o   tabe -ta   no?
       T   -toka-NOM who -ACC eat   -PAST Q
       'What did also Taro eat?'

   b. Nani$_i$ -o   Taro -toka-ga t_i tabe -ta   no?
      who   -ACC T.   -toka-NOM   eat   -PAST Q
      'What did also Taro eat?'

So far, we have observed several similarities between *-toka/tari* and *-mo/sae* in Japanese. In the next section, we turn to some striking differences between them.

2.2 Differences between -toka/tari and -mo/sae

We show that *-toka* and *-tari* obtain disjunctive interpretations in non-veridical contexts (Ladusaw 1996, Giannakidou 2006, 2011, inter alia), though they are conjunctive in regular declarative (veridical) environments as we have observed in the previous section. Such a conjunctive-disjunctive alternation is never possible with *-mo* and *-sae*. They are conjunctive regardless of their environment and carry additive presuppositions that project out of non-veridical contexts, as illustrated in conditionals (10).

(10) a. Taro -mo kita -ra, Ryo -wa   ocha -o   das -u.
        T.   -mo come -if   R.   -TOP   tea -ACC serve -PRES
        'If Taro also comes to the party, Ryo serves tea.'

                                                                        (*disjunctive)

    b. Taro-ga   burokkori -o   tabe -sae   sur -eba, mama   -wa yorokob -u.
       T.   -NOM   broccoli -ACC eat -sae   do -if   mom   -TOP be.happy -PRES
       'If Taro even eats broccoli then his mom will become happy.'

                                                                        (*disjunctive)

   In (11), *-toka* and *-tari* lack such presuppositions. Instead, they gain an inclusive disjunctive interpretation. Thus, in (11a), Ryo serves tea if at least one member, including, but not limited to, overtly pronounced *Taro*, in the relevant set of individuals comes. Likewise, in (11b), Taro's mom becomes happy if Taro does at least one thing reminiscent of a predicate *eating broccoli*. The same applies when *-toka* and *-tari* appear in coordination, demonstrated in (12). Note that *-toka* and *-tari* do not entail that Taro himself comes to the party along with someone else, or that Taro actually eats broccoli, unlike *-mo* 'also' and *-sae* 'even' in (10), which do.

(11) a. Taro -toka-ga   kita -ra,   Ryo -wa ocha -o   das -u
        T.   -toka-NOM come -if   R.   -TOP tea   -ACC serve -PRES
        'If Taro (or someone else) comes to the party, Ryo will serve tea.'

   b. Taro-ga   burokkori -o   tabe -tari-su-reba, mama   -wa yorokob   -u
        T.   -NOM broccoli   -ACC eat -tari-do-if   mom   -TOP be.happy -PRES
        'If Taro eats broccoli (or does something else) his mom will become happy.'

(12) a. Taro-toka Hanako -toka-ga   kita -ra,   Ryo -wa ocha -o   das -u
        T.   -toka H.   -toka-NOM come -if   R.   -TOP tea   -ACC serve -PRES
        'If Taro or Hanako (or someone else) comes to the party, Ryo serves tea.'

   b. Taro-ga   ringo -o   tabe-tari mizu -o   non –dari si-tara mama-wa yorokob-u
        T.   -NOM apple -ACC eat-tari   water -ACC drink-tari do-if   mom -TOP be.happy-PRES
        'If Taro eats an apple or drinks water (or does something else) his mom will be happy.'

This alternation between conjunctive and disjunctive readings is not limited to conditionals as in (11) and (12). For example, in (13), *-toka* and *-tari* are embedded under the possibility epistemic modal *-(rar)e/u* 'can/may' and receive disjunctive readings. Coordinated counterparts are in (14), which obtain disjunctive readings in the same manner as in (13).

(13) a. shoosin   -toka-ga   ari-u/e-ru
        promotion -toka-NOM be-MOD-PRES
        'There is a possibility that you attain promotion or some other thing.'

   b. Godzilla-wa   machi-o   hakaisi-tari si-u-ru
        Godzilla-TOP town-ACC destroy-tari do-MOD-PRES
        'Godzilla may destroy the town or do some other thing.'

(14) a. shoosin   -toka kaigaikimmu   -toka-ga   ari-u/e -ru
        promotion -toka overseas.assignment-toka-NOM be-MOD -PRES
        'There is a possibility that you will get promoted or work abroad (or some other thing).'

   b. Godzilla-wa machi   -o   hakaisi -tari, teki   -o   taosi -tari si -u   -ru
        Godzilla-TOP town   -ACC destroy -tari, enemy   -ACC defeat-tari do-MOD -PRES
        'Godzilla may destroy the town or defeat his enemies (or do some other thing).'

This sharply contrasts with *-mo* and *-sae*, which never gain such disjunctive readings in the same environment. Both examples in (15) must be interpreted conjunctively, as the English translation shows.

(15) a. shoosin   -mo /sae   ari -u/e -ru
       promotion -also /even   be -MOD -PRES
       Lit. 'There is a possibility that you also/even will get a promotion.'

                                                                                    (*disjunctive)

     b. Godzilla-wa   machi-o   hakaisi-sae-si-u-ru
        Godzilla-TOP   town-ACC destroy-sae-do-MOD-PRES
        Lit. 'Godzilla may even destroy the town.'

                                                                                    (*disjunctive)

Yet another environment is inside the scope of other types of epistemic modal particles such as *yooda/rasii/hazuda/kamosirenai* 'maybe/should/seem'. Again, *-toka* and *-tari* obtain a disjunctive interpretation in both (16) and (17), the latter of which involves coordination. In (16a), Taro did not necessarily have beer. The sentence is true if he had had at least some drink similar to beer, which overtly appears. In the same vein in (16b), the professor does not need to smoke cigarettes. Needless to say, the same applies to the coordinated examples in (17).

(16) a. Taro -wa   biiru -toka-o   non -da   {yooda/rasii}/hazuda/kamosirenai
         T. -TOP   beer -toka-ACC   drink -PAST seem/should/maybe
        'Taro seemed to drink/should have/might have drunk beer or something.'

     b. Sensei -wa tabako -o   sut   -tari su-ru   {yooda/rasii} /hazuda/kamosirenai
        prof. -TOP tobacco -ACC smoke   -tari do-PRES seem/should/maybe
        'The professor seems to/should/may smoke tobacco or do something.'

(17) a. Taro -wa   biiru-toka nihonshu-toka-o   non -da {yooda/rasii}/hazuda/kamosirenai
         T. -TOP beer -toka sake   -toka-ACC drink-PAST seem/should/maybe
        'Taro seemed to drink/should have/might have drunk beer or sake (or something).'

     b. Sensei -wa tabako -o   sut   -tari, sake   -o   non -dari su-ru
        prof. -TOP tobacco -ACC smoke   -tari alcohol -ACC drink -tari do-PRES
        {yooda/rasii}/hazuda/kamosirenai
         seem/should/maybe
        'The professor seems to/should/may smoke tobacco or drink alcohol (or do something).'

The relevant variable behavior is absent in the case of *-mo* and *-sae*, as demonstrated in (18).

(18) a. Taro -wa   biiru -mo /sae   non -da   {yooda/rasii}/hazuda/kamosirenai
         T. -TOP   beer -also/even   drink -PAST seem/should/maybe
        Lit. 'Taro seemed to drink/should have/might have drunk also/even beer.'

                                                                                    (*disjunctive)

     b. Sensei -wa tabako -o   sui   -mo/sae su-ru {yooda/rasii}/hazuda/kamosirenai
        prof. -TOP tobacco-ACC   smoke -also/even do-PRES seem/should/maybe

Lit. 'The professor seems to/should/may even/also smoke tobacco.'

(*disjunctive)

Yet another context we will observe is within the scope of a universal quantifier, as illustrated in (19) and (20). Disjunctive readings are only obtained with *-toka* and *-tari*, either when they are used as stand-alone particles in (19) or as coordinators in (20). On the other hand, *-mo* and *-sae* do not, as demonstrated in (21). The sharp contrast between *-toka/tari* and *-mo/sae* leads us to conclude that *-toka* and *-tari* are special among these focus particles.

(19) a. Suugaku -toka-o   bennkyoo-si -ta   gakusei-minna(-ga) siken -ni   ukat-ta
      math   -toka-ACC study   -do -PAST   student-every-NOM exams-DAT pass-PAST
      'Every student that studied math or something passed the exams.'

   b. Gakkai-happyoo-o   si -tari si-ta   innsei   -zen'in-ga   syuusyoku-si-ta
      conf.talk   -ACC do -tari do-PAST graduate.student-all   -NOM get.tenure-do-PAST
      'Every grad student that gave talks in conferences or did something got tenure.'

(20) a. Suugaku   -toka kagaku-toka-o   bennkyoo -si -ta   gakusei -minna (-ga)
      math   -toka science-toka-ACC study   -do-PAST   student -every -NOM
       siken -ni   ukat-ta
       exams -DAT pass-PAST
      'Every student that studied math or science (or something) passed the exams.'

   b. Gakkai-happyoo-o   si -tari, ronbun -keisai -o   si -tari si-ta
      conf.talk   -ACC do -tari paper   -publish -ACC do -tari do-PAST
      innsei   -zen'in -ga   syuusyoku si -ta
      graduate.student -all   -NOM get.tenure   do -PAST
      'Every grad student that gave talks in conferences or published his/her papers (or did something like that) got tenure.'

(21) a. Suugaku-mo /sae bennkyoo-si -ta   gakusei-minna(-ga)   siken   -ni   ukat-ta
      math   -also/even study   -do-PAST student-every -NOM exams   -DAT pass-PAST
       Lit. 'Every student that studied also/even math something passed the exams.'

(*disjunctive)

   b. Gakkai-happyoo-o si -sae si -ta   innsei   -zen'in-ga syuusyoku-si -ta
      conf.talk-ACC   do -even do-PAST graduate.student -all   -NOM get.tenure -do -PAST
       Lit. 'Every grad student that even gave talks in conferences got tenure.'

(*disjunctive)

2.3 Interim conclusion

To summarize, *-toka* and *-tari* exhibit many parallels with the focus particles *-mo* and *-sae* syntactically, but differ from them in their semantic properties. Only *-toka* and *-tari* are sensitive to the (non-)veridicality of their environment and exhibit an alternation between conjunctive and disjunctive readings. In the next section, we propose a semantic analysis to capture our observations of *-toka* and *–tari*, couched in alternative semantics.

3. Analysis
3.1 Syntax of non-exhaustive particles in Japanese

We follow the spirit of previous analyses of *-mo* 'also' (Mitrović and Sauerland 2014), and claim that *-toka* and *-tari* are actually not the coordinator head. Rather, they are focus particles that attach to each conjunct coordinated by a silent coordinator J^0, as in (22). This structure is also illustrated in the tree diagram in (23).

(22)   a. [$_{JP}$ [$_{tokaP}$ Nominal-toka] [$_{J'}$ J [$_{tokaP}$ Nominal-(toka)]]][4]

     b. [$_{JP}$ [$_{tariP}$ Predicate-tari] [$_{J'}$ J [$_{tariP}$ Predicate-tari]]]

(23)
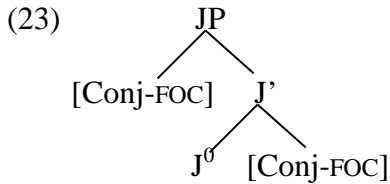
On its own, *-toka* selects for a nominal complement. On the other hand, we propose that *-tari*, given parallels between it and *-mo* and *-sae*, selects a predicative projection below TP.[5]

3.2 Alternative semantics

We propose that sentences with *-toka* and *-tari* simply denote a set of individual and predicate alternatives[6], respectively, as in (24a) and (24b), with no additive presuppositions like those that come with *-mo* and *-sae*.

(24)   a. ⟦Taro-toka⟧ = {Taro, Ryoichiro, Ziro, ...}

     b. ⟦heya-o soojisi-tari⟧
       = {λx.λw.x clean the room, λx.λw.x study English, λx.λw.x eat dinner, ...}

In the case of coordination with *-toka* and *-tari*, we depart from Mitrović and Sauerland's treatment of J^0. We analyze it as simply collecting alternatives introduced by each conjunct in exactly the same way that *or* does in the analysis of Alonso-Ovalle (2006, 2008), as in (25).

---

[4] Since the appearance of the second *-toka* is optional and does not affect the semantics, we assume that it is syntactically always there, but optionally has phonetic content in (22).

[5] At this juncture, we are still agnostic about what exactly *-tari* coordinates or takes as its complement. See Kobayashi & Smith (2017) for further details.

[6] An anonymous reviewer suggests that even in non-coordinative cases a covert conjunction/disjunction is necessary to capture the fact that a conjunctive/disjunctive reading is available in these contexts. However, this is not necessary, as the interaction between the semantics of alternatives generated by *-toka* and *-tari* and the propositional operators we introduce below guarantees the availability of conjunctive and disjunctive readings with no need for covert coordination.

(25) Where ⟦XP⟧ and ⟦YP⟧ ⊆ Dτ, ⟦[[XP] [J [YP]]]⟧ ⊆ Dτ = ⟦XP⟧ ∪ ⟦YP⟧[7]

The alternatives are composed with other elements of the sentence via Pointwise Functional Application (Hamblin 1973), ultimately yielding a set of propositional alternatives.

(26) a. ⟦Taro-toka ga kita⟧
$= \{\lambda w.$Taro came, $\lambda w.$Ryoichiro came, $\lambda w.$Ziro came, ...$\}$

    b. ⟦Taro wa heya-o soojisi-tari sita⟧
$= \{\lambda w.$Taro cleaned the room, $\lambda w.$Taro studied Eng-lish, $\lambda w.$Taro ate dinner, …$\}$

Once the alternatives become propositional, they can be manipulated by one of two propositional quantifiers, defined below (Kratzer and Shimoyama 2002; Alonso-Ovalle 2006, 2008) in (27).

(27) a. $\llbracket \exists \rrbracket^w (A) = \{\lambda w'. \exists p \in A \ \& \ p(w')\}$

    b. $\llbracket \forall \rrbracket^w (A) = \{\lambda w'. \forall p \in A \rightarrow p(w')\}$

In non-veridical contexts, the set of alternatives is existentially quantified as in (27a), which gives rise to the interpretation that at least one of the propositions in the alternative set is true, but not necessarily the one overtly mentioned. The relevant example is repeated here as in (28).

(28) a. Taro -toka (Hanako -toka)-ga   kita -ra,   Ryo -wa ocha -o   das -u
      T.   -toka H.   -toka -NOM come -if   R.   -TOP tea   -ACC serve -PRES
      'If Taro (or Hanako or someone else) comes to the party, Ryo serves tea.'

    b. Taro-ga   ringo-o   tabe-tari mizu -o   non -dari si-tara mama-wa yorokob-u
      T.   -NOM apple-ACC eat -tari water-ACC drink-tari do-if   mom-TOP be.happy-PRES
      'If Taro eats broccoli (or drinks milk or does something else) his mom becomes happy.'

Here we derive the relevant reading by existentially quantifying the set of alternatives in the antecedent of the conditional, resulting in the interpretation of (28a) as in (29). This will ensure that Yosuke will serve tea just as long as at least one of the propositions in the alternative set is true.

(29) {(∃p ∈ {Taro comes in w, Hanako comes in w, …}& p(w)) → Yosuke serves tea in w}

In veridical contexts, the alternatives are instead universally quantified with the operator in (27b), which makes all of the propositions in the alternative set true, and thus gives rise to the conjunctive interpretation observed in (1) and (2), repeated here as (30).

(30) a. Taro -toka   (Hanako -toka)-ga   ki   -ta

---

[7] This allows the alternatives to be composed with other elements of the sentence in the same way regardless of whether coordination is present or not. We refer readers to Smith and Kobayashi (in prep.) for detailed explanations of how the analysis connects with Alonso-Ovalle's (2006, 2008) proposal for *or*.

    T.   -toka   H.   -toka-NOM   come -PAST
    'Taro, (Hanako,) and others came.'
  b. Taro -wa heya -o   soojisi -tari (eigo   -o   benkyoosi -tari) si -ta
    T.   -TOP room -ACC clean   -tari English -ACC study   -tari do -PAST
    'Taro cleaned his room, (studied English,) and did other things.'

Here, the universal quantifier is inserted, resulting in an interpretation for (30a) as in (31).

(31) {∀p ∈ {Taro came in w, Hanako came in w, …} → p(w)}

Here, the alternative set has been universally quantified, and (31) states that every proposition in the set of alternatives holds in the world of evaluation. This brings about what we have termed the conjunctive interpretation.

4. Conclusion

In this paper, we have shown that *-toka* and *-tari* pattern much like the focus particles *-mo* and *-sae* in terms of their syntactic distribution: *-toka* and *-tari* can be used as stand-alone particles and as polysyndetic coordinators, are incompatible with topical *-wa*, and induce intervention effects, etc. However, they differ from other focus particles in lacking additive presuppositions and having interpretations sensitive to the (non-)veridicality of their environment.

    To sum up, this paper proposes an analysis of these particles as introducing alternatives, which are then manipulated by propositional quantifiers higher in the structure, and unifies their use as single particles with their use as coordinators.

    Before we conclude, we briefly touch on some potential shortcomings. Our analysis falls short of explaining why *-toka* and *-tari* in veridical contexts are exclusively associated with ∀. At the same time, it fails to capture why they are associated with ∃ in non-veridical contexts. *Taro-toka-ga kita* 'Taro-toka came' in veridical contexts definitely means that 'at least Taro came', but the definition of the operators in (27), repeated here as (30), alone does not guarantee this fact.

(29)   a. $[\![∃]\!]^w (A) = \{λw'. ∃p ∈ A \& p(w')\}$

     b. $[\![∀]\!]^w (A) = \{λw'. ∀p ∈ A → p(w')\}$

    One possibility is that the quantifiers are not inserted freely, but are inserted on the basis of relative semantic strength depending on the context. For instance, the universal quantifier is stronger in the veridical contexts, while the existential quantifier is more appropriate to the non-veridical contexts. Another possible approach is to rely less on the existential quantifier and to let the different semantic environments carry most of the burden. For instance, Alonso-Ovalle (2006) proposes that conditionals universally quantify over sets of alternatives introduced by disjunction. This approach would work for conditionals with *–toka/-tari* as in (11): (11a) would receive a meaning in which all closest worlds where Taro comes are worlds where Ryo serves tea, all closest worlds where Hanako comes are those where Ryo serves tea, etc. Crucially, this would not require both Taro and Hanako to come in the same world, the desired result. Whether this general analytical strategy would work generally is a topic for future research.

Another point is concerning the overgeneration problem. Let us take a simple sentence *Taro-toka-ga ki-ta* 'Taro-toka came' as an example. Our analysis is a bit too strong in that the system allows *Taro-toka* to denote all the individuals in the contextually salient set. However, *Taro-toka-ga ki-ta* 'Taro-toka came' is true of at least one other contextually salient individual alternative, but not necessarily all of them. One way to deal with this is to further restrict the alternative set using an explicit similarity relation, where those individuals/predicates that are included in the alternative set are just those that are similar to the overtly mentioned arguments of *–toka/-tari* in some contextually defined sense.

In future work, we intend to pursue some of these tentative suggestions a bit further. See Smith & Kobayashi (in prep.) for further details.

Ryoichiro Kobayashi[1,2] and Ryan Walter Smith[3]
[1] Graduate School of Languages and Linguistics, Sophia University[1]
[2] Japan Society for the Promotion of Science[2]
[3] University of Arizona
ryoichiro.k@eagle.sophia.ac.jp (Kobayashi)
ryanwaltersmith@email.arizona.edu (Smith)

Abbreviations

NOM	nominative marker	PAST	past tense
ACC	accusative marker	MOD	modal head
DAT	dative marker	FOC	focus particle
TOP	topic marker	Q	question marker
PRES	present tense	Conj	conjunct

References

Alexiadou, A., Iordăchioaia, G., and F. Schäfer. (2011). Scaling the variation in Romance and Germanic nominalizations. In P. Sleeman and H. Perridon (eds.), *Variation and Change in the Romance and Germanic Noun Phrase*, 25-40. Amsterdam: John Benjamins

Alonso-Ovalle, L. (2006). Disjunction in alternative semantics. Doctoral dissertation, University of Massachusetts Amherst.

Alonso-Ovalle, L. (2008). Innocent exclusion in an alternative semantics. *Natural Language Semantics*, 16(2), 115-128.

Giannakidou, A. (2006). Only, emotive factives, and the dual nature of polarity dependency. *Language*, 82, 575-603.

Giannakidou, A. (2011). Negative and positive polarity items. In K. von Heusinger, C. Maienborn & P. Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning 2*, 1660-1712. Berlin: De Gruyter Mouton.

Hamblin, C. L. (1973). Questions in Montague English. *Foundations of language*, 10(1), 41-53.

Hoji, H. (1986). Scope interpretation in Japanese and its theoretical implications. In *WCCFL*, 5, 87-101.

Kageyama, T. (1993). *Bunpou to gokeisei* [*Grammar and word formation*]. Tokyo: Hituzi Shobo.

Kobayashi, R. (2017). Is it syntactic or pragmatic? A hybrid analysis for LF-intervention effects. *Language and Linguistics*, 18(2), 228-253.

Kobayashi, R. & Smith, R.W. (2017). A case of semantic restriction on coordination. In *Proceedings of SICOGG 19*, 451-463. Seoul National University.

Kratzer, A., and J. Shimoyama. (2002). Indeterminate pronouns: The view from Japanese. In *Proceedings of the 3rd Tokyo conference on psycholinguistics* 1-34. Tokyo: Hitsuji Shobo.

Ladusaw, W. (1996). Negation and polarity items. In S. Lappin (ed.), *Handbook of Contemporary Semantic Theory*, 299-319. Oxford: Blackwell.

Mitrović, M., and U. Sauerland. (2014). Decomposing coordination. *NELS*, 44, 39-52.

Smith, R. W., and R. Kobayashi. (in prep.). The syntax and semantics of non-exhaustive coordination in Japanese. [MS], University of Arizona and Sophia University.

Tomioka, S. (2007). Pragmatics of LF intervention effects: Japanese and Korean wh-interrogatives. *Journal of Pragmatics*, 39(9), 1570-1590.

The presupposition of exclamatives at the syntax-semantics interface: Evidence from German and Japanese

Katsumasa Ito

German *wh*-exclamatives allow verb-last (VL) word order as well as verb-second (V2) word order when they contain degree expressions. Japanese *koto*-exclamatives are ill-formed when their predicates are not gradable. The goal of this paper is to show that these properties of German exclamatives and Japanese exclamatives can be explained when we analyze the degree components of exclamative clauses as not presupposed. In order to implement this idea, I adopt Zanuttini & Portner's (2003) proposal that the presupposition trigger of exclamatives is a factivity-operator (FACT-Op) at the left periphery of exclamative clauses. The conclusion of this paper supports Zanuttini & Portner's (2003) primary intention that the presupposition of exclamatives can be captured (not only at the semantic level but also) at the syntactic level.

1. Introduction

German is a so-called V2 (verb-second) language and the finite verb of a main clause is placed in the second position, while the finite verb of a subordinate clause is placed in the final position. However, in German *wh*-exclamatives, the placement of the finite verbs is a little tricky. The examples in (1) and (2) show that *wie* 'how' exclamatives (1a, b) and *was-für-ein* 'what a' exclamatives (2a, b) allow VL (verb-last) order as well as V2 order.

(1) a. Wie schön   die getanzt **hat**!
        how beautifully she danced has
        'How beautifully she danced!'
    b. Wie schön   **hat** die getanzt!
        how beautifully has she danced
        'How beautifully she danced!'

(2) a. Was für ein tolles Auto der gekauft **hat**!
        what for a nice car he bought has
        'What a nice car he bought!'

b. Was für ein tolles Auto **hat** der gekauft!
    what for a   nice car   has he bought
    'What a nice car he bought!'

However, not all German *wh*-exclamatives allow V2 order (cf. Rosengren 1992). As shown in (3) and (4), some *wh*-exclamatives allow only VL order.

(3) a. Wen ich getroffen **habe**!
         who I   met   have
         'The person who I met!'
     b. *Wen **habe** ich getroffen!
          who have I   met

(4) a. Wo   der gewesen **ist**!
         where he been   is
         'The place he has been to!'
     b. *Wo   **ist** der gewesen!
          where is   he   been

The first question I would like to raise in this paper is why some German *wh*-exclamatives allow V2 order as well as VL order, while others allow only VL order. It is shown that the exclamatives that allow V2 order assert degree meanings and that this assertion licenses V2.[1] The implementation of this analysis becomes possible when we adopt Zanuttini & Portner's (2003) proposal that the presupposition trigger of exclamatives is a FACT-Op (factivity-operator) at the left periphery of exclamative clauses.

  This paper also investigates Japanese *koto*-exclamatives. *Koto* is a nominalizer in Japanese. While nominalized clauses can be embedded as shown in (5), they can be interpreted as exclamatives when they are independent (6). Note that the subject is marked genitive in (6).

(5) Yuki-wa [Sachiko-ga   kirei   datta koto] -o   shitteiru.
     Yuki-TOP Sachiko-NOM beautiful was   NMLZ -ACC know
     'Yuki knows that Sachiko was beautiful.'
(6) Sachiko-no   kirei   datta koto!
     Sachiko-GEN beautiful was   NMLZ
     'How beautiful Sachiko was!'

The *koto*-exclamatives have a restriction on their predicates. If the predicate of *koto*-exclamatives is not gradable, the sentence is unacceptable. In (7), the predicate is *kekkonshita* 'got married,' which is not gradable.

(7) * Sachiko-no   kekkonshita koto!
       Sachiko-GEN got married NMLZ
       Intended: 'Sachiko got married!'

---

[1] Note that this does not mean only V2 order has assertional degree meaning. Assertion is a necessary but not a sufficient condition for V2 order (see Antomo & Steinbach 2010 and Reis 2013).

The second question of this paper is why there is such a restriction on Japanese *koto*-exclamatives. I argue that *koto*-exclamatives cannot host FACT-Op because of the absence of C (cf. Miyagawa 2011) and that this is the reason for the restriction on the predicates.

At first sight, the two phenomena, namely the verb positions in German and the restriction on predicates in Japanese, seem to be unrelated. This paper shows, however, that the both phenomena can be explained when we assume that the degree components of exclamative clauses are not presupposed. In order to implement this idea in a formal way, I adopt Zanuttini & Portner's (2003) proposal that the presupposition trigger of exclamatives is FACT-Op.

This paper is organized as follows: Section 2 observes the behavior of German *wh*-exclamatives and shows that their degree components are asserted, while the rest of the proposition is presupposed. It is shown that V2 order is allowed when the exclamative clause contains a degree expression and that assertion of the degree meaning licenses the V2 order of German *wh*-exclamatives. Section 3 introduces FACT-Op, which plays a role of a presupposition trigger of exclamatives. In order to justify the presence of FACT-Op at the syntactic level, factive clauses in German and Japanese are investigated. Section 4 shows how the degree components outscope FACT-Op and are asserted in German. Following Corver (1990) and Rosengren (1992), I assume that *wh*-phrases in exclamatives constitute a Degree Phrase (DegP). It is argued that the Agree-relation between CP and DegP enables degree component to outscope FACT-Op. In Section 5, the reason why there is a restriction on the predicates of Japanese *koto*-exclamatives is presented. I adopt Miyagawa's (2011) proposal with regard to genitive subjects in Japanese and assume that there is no C in *koto*-exclamatives. Because of the absence of C, *koto*-exclamatives cannot host FACT-Op and this accounts for the restriction on predicates. Section 6 comprises the conclusion.

2. Presuppositions of German wh-exclamatives

In this section, I will run three tests to show that the degree meanings of exclamatives are not presupposed. A lot of previous studies suggest that the propositions of exclamatives are presuppositional (cf. Grimshaw 1979, Zanuttini & Portner 2003, Abels 2010, Driemel 2015). While most of these studies must rely on the assumption that an exclamative can be embedded, Driemel (2015) proposes a novel test which can prove the presuppositional status of matrix exclamatives in German. She observes that VERUM-focus[2] (cf. Höhle 1992, Romero & Han 2004) of a presupposed clause is exceptionally licensed under *denial contexts*[3] (i.e. in the case of common ground revision) as in (8). Otherwise VERUM-focus of a presupposed clause is infelicitous.

(8) A: Hanna likes company when she visits the opera, which is why she is angry about the
         fact that her daughter did not go with her this time.
     B: You're wrong – Hanna likes it most when she goes alone.
         Sie ärgert   sich   darüber, dass ihre Tochter gestern   mit ihr in der Oper WAR.
         she is.angry herself about   that her daughter yesterday with her in the opera was
         'She is angry that her daughter DID accompany her.'

<div align="right">(Driemel 2015:422)</div>

---

[2] The function of VERUM-focus is to emphasize or insist on the truth or falsity of the proposition. In German, a focus accent on an auxiliary or a copular verb leads to a VERUM effect (cf. Höhle 1992).
[3] The context is *denial* when the truth of the proposition is explicitly denied.

The same behavior is observed in German *wh*-exclamatives. *Denial contexts* like (9) exceptionally licenses *wh*-exclamatives with VERUM-focus, though German *wh*-exclamatives are generally incompatible with VERUM-focus.

(9) A: Peter is not a big traveler. The places he has not been to!
    B: Aber wo   der auch schon gewesen IST!
       but   where he also already been   is
      'But the places he HAS been to already!'
      Just think about the time before he started studying.

<div align="right">(Driemel 2015: 423)</div>

Adopting Rooth's (1992) alternative semantics, Driemel (2015) presents an elegant formal account for this behavior of VERUM-focus. Space restrictions do not permit a description of the entire sketch of her theory, but now we know one remarkable property of presupposed matrix clauses: they are compatible with VERUM-focus in denial contexts.

   Driemel (2015) does not discuss the presuppositional status of German *wh*-exclamatives that allow V2 order like (1) and (2). When we apply her VERUM-focus test to (1) and (2) and try to deny the degree meaning of the utterance, we get interesting results. As shown in (10) and (11), the exclamatives are infelicitous in denial contexts.

(10) A: I know Mary is a good dancer, but today her dance was not so good.
     B: #Wie schön   die getanzt HAT!
        how beautifully she danced has
        'How beautifully she DID dance!'

(11) A: Peter bought a new car, but I thought the car was not so nice.
     B: #Was für ein tolles Auto der gekauft HAT!
        what for a   nice car   he bought has
        'What a nice car he DID buy!'

In (10) and (11), the exclamatives cannot deny the degree meaning of A's utterance. I have no clear explanation for these data, but at least it seems safe to say that this pattern suggests that the degree meaning of exclamatives is not presuppositional: if the degree meaning of the exclamatives were presuppositional, they should be compatible with VERUM-focus in denial contexts, which is contrary to the fact.

   Von Fintel's (2004) *Hey-wait-a-minute*-test[4] also suggests that the degree meaning of exclamatives is not presuppositional. As shown in (12B') and (13B'), it is odd for hearer B to claim that he/she did not know the degree meaning after saying *warte mal* 'wait a minute'.

---

[4] Von Fintel (2004) presents only English data. In German, hearer B's claim that he/she did not know the presupposed content can be preceded by *warte mal* 'wait a minute' (iB), while the asserted content cannot (iB').
(i) A: Peter bereut, dass er den Kuchen gegessen hat.
     Peter regrets that he the cake   eaten   has
     'Peter regrets that he ate the cake'
  B: Warte mal!   Ich wusste nicht, dass er den Kuchen gegessen hat.
     wait a minute I   knew   not   that he the cake   eaten   has
     'Hey, wait a minute! I have no idea that he ate the cake.'
  B': #Warte mal!   Ich wusste nicht, dass er es bereut.

(12) A: Wie schön   die getanzt hat!
       how beautifully she danced has
       'How beautifully she danced!'
  B: Warte mal!   Ich wusste nicht, dass sie getanzt hat.
      wait a minute I   knew not   that she danced has
      'Hey, wait a minute! I didn't know that she danced.'
  B': ?#Warte mal!   Ich wusste nicht, dass es so schön   war.
        wait a minute I   knew not   that it so beautiful was
        'Hey, wait a minute! I didn't know that it was so beautiful.'

(13) A: Was für ein tolles Auto der gekauft hat!
       what for a   nice car he bought has
       'What a nice car he bought!'
  B: Warte mal!   Ich wusste nicht, dass er ein Auto gekauft hat.
      wait a minute I   knew not   that he a car   bought has
      'Hey, wait a minute! I didn't know that he bought a car.'
  B': ?#Warte mal!   Ich wusste nicht, dass es so toll ist.
        wait a minute I   knew not   that it so nice is
        'Hey, wait a minute! I didn't know that it was so nice.'

Note that the proposition without the degree meaning can be complained about as shown in (12B) and (13B). This suggests that the non-degree components of the propositions of the exclamative are presupposed.

   Another suggestive test is *No*-test (cf. Chernilovskaya et al. 2012).[5] As shown in (14), the hearer B can negate the asserted content by uttering *no* (14B), while the presupposed content cannot be negated after *no* (14B').

(14) A: Peter bereut, dass er den Kuchen gegessen hat.
       Peter regrets that he the cake   eaten   has
       'Peter regrets that he ate the cake'
  B: Nein, er bereut es nicht.
      no   he regrets it not
      'No, he does not regret it.'
  B': ?#Nein, er hat den Kuchen nicht gegessen.
        no   he has the cake   not   eaten
        'No, he didn't eat the cake.'

The results of *No*-test on German *wh*-exclamatives are presented in (15) and (16). These data also suggest that the degree meanings of exclamatives are not presupposed.

(15) A: Wie schön   die getanzt hat!
       how beautifully she danced has
       'How beautifully she dance!'

---

        wait a minute I   knew not   that he it regrets
         'Hey, wait a minute! I have no idea that he regrets it.'

B: Nein, das war nicht so schön.
    no   that was not   so beautiful
    'No, it was not so beautiful.'

B': ?#Nein, sie hat nicht getanzt.
    no   she has not   danced
    'No, she didn't dance.'

(16) A: Was für ein tolles Auto der gekauft hat!
        what for a   nice   car   he   bought has
        'What a nice car he bought!'
    B: Nein, es war nicht so toll.
        no   that was not   so nice
        'No, it was not so nice.'
    B': ?#Nein, er hat kein Auto gekauft.
        no   he has no   car   bought
        'No, he bought no car.'

As for the German *wh*-exclamatives that do not allow V2 order such as (3) and (4), the propositions seem to be presupposed, as (17) and (18) show.

(17) A: Wen Peter getroffen hat!
        who Peter met   has
        'The person who Peter met!'
    B: Warte mal! Ich wusste nicht, dass er jemanden getroffen hat.
        wait a minute I   knew   not   that he somebody met   has
        'Hey, wait a minute! I didn't know that he met somebody.'
    B': ?#Nein, er hat niemanden getroffen.
        no   he has nobody   met
        'No, he met nobody.'

(18) A: Wo   Peter gegangen ist!
        where Peter gone   is
        'The place Peter went to!'
    B: Warte mal! Ich wusste nicht, dass er irgendwo   gegangen ist.
        wait a minute I   knew   not   that he somewhere gone   is
        'Hey, wait a minute! I didn't know that he met somebody.'
    B: ?#Nein, er ist nirgends gegangen.
        no   he is   nowhere gone
        'No, he has gone nowhere.'

When we adopt the idea that the degree meanings of German *wh*-exclamatives are asserted and the non-degree components of the propositions are presupposed, the reason why exclamatives like (1) and (2) allow V2 order can be explained. In German, V2 order is allowed if the proposition has assertion status (cf. Lohnstein 2000, Meinunger 2004, Truckenbrodt 2006, Antomo & Steinbach 2010, Reis 2013). Typical examples are presented in (19) and (20).

(19) a. Ich glaube, dass er Recht **hat**.
      I   think   that he right   has
      'I think that he is right.'

   b. Ich glaube, er **hat** Recht.
      I   think   he has right
      'I think he is right.'

(20) a. Ich bereue, dass ich den Kuchen gegessen **habe**.
      I   regret that I   the cake   eaten   have
      'I regret that I ate the cake.'
   b. *Ich bereue, ich **habe** den Kuchen gegessen.
       I   regret I   have the cake   eaten

In (19), the embedded clause has assertion status because of the matrix predicate *glauben* 'think' and the V2 order of the embedded clause is licensed (19b). On the other hand, V2 order of the embedded clause is banned in (20) because the clause is presupposed due to the matrix predicate *bereuen* 'regret' (20b).

   The exclamatives like (1) and (2) have degree meaning, while (3) and (4) do not. This is supported by the fact that (3) and (4) can be uttered under contexts without degree, see (21) and (22).[6]

(21) Context:
    Speaker A visited a conference and expected to meet scholars B, C and D. A didn't expect to meet scholar E. However, contrary to A's expectation, A ended up meeting E.
    A: Wen ich getroffen habe!
       who I   met   have
       'The person who I met!'
    A': #Was für einen Forscher ich getroffen habe!
         what for a   scholar I   met   have
         'What a scholar I met!'

(22) Context:
    Speaker A believed that Peter had visited Berlin or Leipzig. However, A is now told that Peter visited Tokyo. Speaker A is surprised because it was totally unexpected.
    A: Wo   der gewesen ist!
       where he been   is
       'The place he has been to!'
    A': #Was für eine Stadt der besucht hat!
         What for a   city he visited has
         'What a city he visited!'

Now we can explain why the exclamatives like (1) and (2) are compatible with V2 order: they assert degree meanings and their assertion status licenses V2. In (3) and (4), V2 order is banned, because their propositions are presupposed (see (17) and (18)).

---

[6] The context (21) is adopted from Rett (2011b:23)

This analysis, however, raises a theoretical question: What is the mechanism of the assertion of the degree meanings? In order to answer this question, the next section introduces FACT-Op.

3. FACT-Op

Zanuttini & Portner (2003) propose FACT-Op in order to capture the presuppositional status of exclamatives at the syntax-semantics interface. They observe that exclamative clauses can be embedded under factive predicates and claim that exclamatives and factive complements have FACT-Op, which ensures the presuppositional status of a proposition. According to Zanuttini & Portner (2003), the structure of the left periphery of exclamatives is (23). They assume a multiple CP construction: The higher CP hosts a *wh*-phrase and the FACT-Op is positioned at the specifier of the lower CP.[7]

(23) $[_{CP1}$ *wh*-phrase C_1 $[_{CP2}$ FACT-Op $C_2]]]$

One may wonder whether we should assume FACT-Op at the syntactic level, but a lot of previous studies suggest that the left periphery of factive complements have such an operator (cf. Watanabe 1993, Rossou 2010, Miyagawa 2012, Haegeman 2014). In this section, I present one phenomenon in German and Japanese that supports the presence of FACT-Op. As shown in (24), adjunct-extraction from a factive clause is impossible in German.

(24) *Wann$_i$ bereust du, dass du   t$_i$ den Kuchen gegessen hast?
        when regret you   that you   the cake   eaten   has

The ill-formedness of (24) is due to the minimality effect. Since FACT-Op is at Spec-CP of the factive complement, the operator intervenes in the extraction.[8, 9] A pure semantic account seems to be difficult, because in Japanese, a *wh*-in-situ language, a corresponding sentence (25) is perfectly acceptable. Our syntactic account correctly predicts that (25) is grammatical: because there is no *wh*-movement, no minimality violation occurs.

(25) Kimi-wa [keeki-o   itsu   tabeta koto] -o   koukaisiteiru no?
        you-TOP   cake-ACC when ate   NMLZ -ACC regret   Q
        'You regret that you ate the cake at a certain time, but what time is that?'

---

[7] According to Zanuttini & Portner (2003:61), Paduan (an Italian dialect spoken in north-eastern Italy) exclamatives have an overt complimentizer. The structure (23), in which the *wh*-phrase is in the higher projection, allows the presence of the overt complimentizer at C_2 in (23) without a doubly-filled-COMP filter violation.

[8] Following Haegeman (2014), I assume that the operator has the same feature with *wh*-phrase. One might argue that this assumption is incompatible with the structure (23), because in (23) the *wh*-phrase crosses FACT-Op, violating minimality. This problem can be avoided, when we assume that the movement of exclamative *wh*-phrase is driven by degree-feature, not by *wh*-feature. I return to this point in section 4.

[9] An anonymous reviewer kindly pointed out that the verb positions of German exclamatives can be explained if one assumes that FACT-Op is located at C head, since V2 order is realized by V-T-C head movement. This analysis, however, is unable to explain why factive clauses like (24) resists adjunct-extraction. If FACT-Op were at C, its specifier position would be available for the extraction, which is contrary to the fact.

   This section introduced FACT-Op of exclamative clauses, which was proposed by Zanuttini & Portner (2003). In the next section, the mechanism of the assertion of the degree component is proposed based on the structure in (23).

4. Assertion of degree meanings in German wh-exclamatives

In order to calculate the degree meaning of exclamatives, I assume that an exclamative *wh*-phrase constitutes DegP (Degree Phrase, cf. Corver 1990, Rosengren 1992, Rett 2008). Following Rett (2008:615), I further assume that DegP can take not only AP but also DP as its complement. Though Rett (2008) does not discuss its empirical advantages on syntactic side, the morphology of German degree word *solch* 'such' supports this assumption. As shown in (26), *solch* does not inflect when it is positioned left to the determiner.

(26) a. solch ein Buch
          such a   book
     b. ein solch**es**   Buch
        a   such.SG.N book

When we assume that the phrases in (26) have the structures in (27), this behavior is explained. In (27a), DP-Phase (cf. Chomsky 2000, Bošković 2005) blocks the agreement of *solch* 'such' with *Buch* 'book'. This is illustrated in (28).

(27) a. [DegP solch [DP ein [NP Buch]]]
     b. [DP ein [NP [DegP solches [AP ∅]] Buch]]

(28)
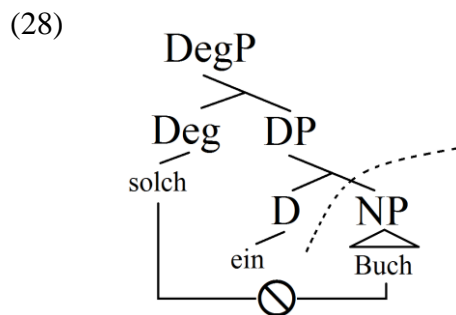
Note that *solch* 'such' ensures the degree meaning. An exclamative clause with *solch* 'such' like (29) cannot be felicitously uttered in the context without degree such as (22).

(29) Dass Peter solch eine Stadt besucht hat!
     that Peter such a   city visited has
     'That Peter visited such a city!'

The assumption that an exclamative *wh*-phrase constitutes DegP is supported by the fact that the exclamative *wh*-phrase cannot co-occur with *solch* 'such.' The presence of a DegP prevents another DegP from appearing.

(30) *Was für solch eine Stadt Peter besucht hat!
       what for such a   city Peter visited has
       Intended: 'What a city Peter visited!'

As semantics of *wie* 'how' and *was für* 'what a' that constitute DegP, I propose (31). *Wie* takes an adverb A, an abstracted proposition f and a contextually-valued degree d as its arguments and yields a truth-value 1 iff f(A) is true and the degree of A exceeds d. *Was für* takes an individual x, an abstracted proposition P and a contextually-valued degree d as its arguments and yields a truth-value 1 iff P(x) is true and the degree of x exceeds d. The function μ maps A or x to a unique degree that represents the measure of A or x with respect to a certain demension (cf. Rett 2011ab, Morzycki 2012, Solt 2015).

(31) a. $\llbracket$wie$\rrbracket$ $= \lambda A_{<et,et>} \lambda f_{<<et,et>,t>} \lambda d_d . \mu(A) \geq d \wedge f(A)$
     b. $\llbracket$was für$\rrbracket$ $= \lambda x_e \lambda P_{et} \lambda d_d . \mu(x) \geq d \wedge P(x)$

Zanuttini & Portner (2003) do not present a formal semantic representation of FACT-Op. In this paper, I tentatively assume that the semantics of FACT-Op is (32). PRESUPP in (32) is a set of presupposed propositions.

(32) $\llbracket$FACT-Op$\rrbracket$ $= \lambda g_t : g \in$ PRESUPP. g

When we adopt (23) for the left periphery of exclamatives, the syntactic structure of (1) is (33).

(33) [CP [DegP wie schön]$_i$ [CP FACT-Op [IP die t$_i$ getanzt hat]]]
            how beautifully   she   danced has

In (33), *wie schön* can cross FACT-Op without violating minimality, because this movement is driven by degree-feature, not by *wh*-feature. The presence of DegP suggests that this analysis is on the right track.

Under these assumptions, the semantic composition of (33) proceeds as in (34).[10] Following Rett (2011ab), I assume that the degree argument d undergoes existential closure. As shown in (34), we get the desired result in which the degree meaning is asserted, while other components of the proposition are presupposed. Note that the right conjunct of the truth-condition *beautifully* (*danced*) (*she*) is trivially fulfilled because of the presupposition.

(34) $\llbracket$wie schön$\rrbracket$ ($\llbracket$FACT-Op$\rrbracket$ ($\llbracket$die *t* getanzt hat$\rrbracket$))
     $=$ $\llbracket$wie schön$\rrbracket$ (*she danced* $\in$ PRESUPP. *she danced*)   **by FA**
     $\Rightarrow$ $\llbracket$wie schön$\rrbracket$ ($\lambda A_{<et,et>}$: *she danced* $\in$ PRESUPP. A (*danced*) (*she*))   **by PA**
     $=$ $\llbracket$wie$\rrbracket$ ($\llbracket$schön$\rrbracket$) ($\lambda A_{<et,et>}$: *she danced* $\in$ PRESUPP. A (*danced*) (*she*))
     $=$ λd_d : *she danced* $\in$ PRESUPP. μ (*beautifully*) $\geq$ d $\wedge$ *beautifully* (*danced*) (*she*) **by FA**

As for the exclamative with *was für* like (2), its syntactic structure is (35) and the semantic composition proceeds as in (36). This analysis also gives the desired result in which the

---

[10] FA and PA in (34) and (36) stand for functional application and predicate abstraction respectively.

degree meaning of German *wh*-exclamative is not presupposed. The right conjunct of the truth-condition *he bought a car* is trivially fulfilled because of the presupposition.

(35) [CP [DegP was   für ein Auto]*i* [CP FACT-Op [IP der t*i* gekauft hat]]]
         what for a   car   he   bought has

(36) ⟦was für ein Auto⟧ (⟦FACT-Op⟧ (⟦der *t* gekauft hat⟧))
   =   ⟦was für ein Auto⟧ (*he bought t* $\in$ PRESUPP. *he bought t*)   **by FA**
   $\Rightarrow$ ⟦was für ein Auto⟧) (λx_e : *he bought* x $\in$ PRESUPP. *he bought* x)   **by PA**
   =   ⟦was für⟧ (⟦ein Auto⟧) (λx_e : *he bought* x $\in$ PRESUPP. *he bought* x)
   = λd_d : *he bought a car* $\in$ PRESUPP. μ (*a car*) $\geq$ d $\land$ *he bought a car*   **by FA**

This section presented the mechanism of how the degree meaning is asserted. The DegP agrees with CP to outscope FACT-Op. This movement does not violate minimality since it is driven by degree-feature. The denotations of *wie* 'how' and *was für* 'what a' were proposed and the semantic compositions of German *wh*-exclamatives were shown. The next section investigates Japanese *koto*-exclamatives, which can be also explained when we assume FACT-Op.

5. Restrictions on the predicate of Japanese koto-exclamatives

In Section 1, we observed the restriction on the predicate of Japanese *koto*-exclamatives. For ease of reference, I repeat the examples (6) and (7) as (37) and (38) below.

(37) Sachiko-no   kirei   datta koto!
     Sachiko-GEN beautiful was   NMLZ
     'How beautiful Sachiko was!'

(38) *Sachiko-no   kekkonshita koto!
      Sachiko-GEN got married NMLZ
      Intended: 'Sachiko got married!'

Japanese *koto*-exclamatives are incompatible with non-degree predicates (38). Note that the subject is marked as genitive in (37). If the subject is marked as nominative, it is difficult to interpret the sentence as exclamative. In embedded clauses, on the other hand, the subject can be marked as nominative (39).

(39) Yuki-wa [Sachiko-ga   kirei   datta koto] -o   shitteiru.
     Yuki-TOP Sachiko-NOM beautiful was   NMLZ -ACC know
     'Yuki knows that Sachiko was beautiful.'

Japanese genitive subjects are licensed typically in embedded contexts. In (40), the subject can be marked as genitive as well as nominative.

(40) [Sachiko {-ga / -no} yonda] hon
      Sachiko -NOM -GEN read   book

'the book which Sachiko read'

According to Miyagawa (2011), the structure of clauses with genitive subjects is (41). It lacks C-projection and thus there is no feature inheritance from C to T (cf. Chomsky 2008). As a result, T does not have formal features and D checks genitive Case of the subject.

(41)
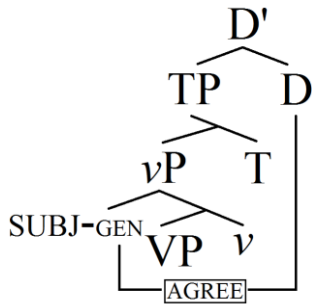
When we adopt Miyagawa's analysis, Japanese *koto*-exclamatives do not have C-projection, since its subject is marked as genitive. This means that *koto*-exclamatives cannot host FACT-Op, because the FACT-Op is positioned at Spec-CP (cf. section 4).

   According to Zanuttini & Portner (2003), exclamatives are characterized by the following two properties.

(42) a. Exclamatives contain a WH operator-variable structure.
      b. Exclamatives contain an abstract morpheme FACT in the CP domain.

                                                       (Zanuttini & Portner 2003:40)

Japanese *koto*-exclamatives seem to conform to the condition (42a). They can contain *wh*-phrase *nante* 'how' (43), though it does not have to be realized phonologically as in (37).

(43) Sachiko-no   nante kirei   datta koto!
      Sachiko-GEN how   beautiful was   NMLZ
      'How beautiful Sachiko was!'

Note that the *wh*-phrase in (43) does not presume the existence of CP, because the *wh*-phrase does not have an uninterpretable feature that must be deleted by C. A *wh*-phrase has an interpretable feature that can delete its counterpart at C (cf. Chomsky 2000). On the other hand, the condition (42b) is not fulfilled by *koto*-exclamatives. They have no abstract morpheme FACT (i.e. FACT-Op) due to the absence of C. This is the reason why (38) cannot be interpreted as exclamatives: since the *koto*-clause of (38) does not have FACT-Op, it is unable to be exclamative. The clause (38) itself is not ill-formed, because it is grammatical when embedded, as shown in (44). The *koto*-clause of (38) just cannot get the interpretation as exclamative.

(44) Yuki-wa [Sachiko-no   kekkonshita koto] -o   shitteiru.
      Yuki-TOP Sachiko-GEN got married NMLZ -ACC know
      'Yuki knows that Sachiko got married.'

The analysis is supported by the fact that an embedded *koto*-clause is incompatible with a *wh*-phrase *nante* 'how', as (45) illustrates. This suggests that the embedded *koto*-clause is not interpreted as exclamative.

(45) ??Yuki-wa [Sachiko-no   nante kirei   datta koto] -o   shitteiru.
　　　 Yuki-TOP Sachiko-GEN how   beautiful was   NMLZ -ACC know
　　　 Intended: 'Sachiko knows how beautiful the scenery was.'

Then, why is the matrix *koto*-clause with a degree predicate felicitous (38)? As shown in section 4, the degree meaning of exclamatives is asserted. (38) just asserts the degree meaning and the proposition is not presupposed. Japanese *koto*-exclamatives are not "exclamative" in the sense of Zanuttini & Portner (2003), because they do not fulfill the condition (42b). *Koto*-exclamatives just assert degree meanings and this is the reason why (38) is felicitous. I speculate here that the degree predicates have assertion status inherently and this enables *koto*-clauses to be used as a matrix clause.

　　The results of the *Hey-wait-a-minute*-test (46B) and *No*-test (46B') suggest that the analysis is on the right track. The proposition of a *koto*-exclamative seems to be asserted as shown in these examples.

(46) A: Sachiko-no   kirei   datta koto!
　　　　　Sachiko-GEN beautiful was   NMLZ
　　　　　'How beautiful Sachiko was!'
　　　B: ?#Chotto matte! [Sachiko-ga   kirei   datta] towa shiranakatta.
　　　　　　little wait   Sachiko-NOM beautiful was   QUOT knew.NEG
　　　　　　'Hey, wait a minute! I didn't know that Sachiko was so beautiful.'
　　　B': Iya, Sachiko-wa sonna kirei   jyanakatta.
　　　　　No Sachiko-TOP such   beautiful was.NEG
　　　　　'No, Sachiko was not so beautiful.'

6. Conclusion

This paper investigated German *wh*-exclamatives and Japanese *koto*-exclamatives. It is shown that their curious properties, namely the verb positions in German and the restriction on predicates in Japanese, can be explained when we assume the existence of FACT-Op. Some German *wh*-exclamatives allow V2 order, because their degree meanings outscope FACT-Op to be asserted. Japanese *koto*-exclamatives are incompatible with non-degree predicates, because the absence of FACT-Op prevents *koto*-exclamatives from being "exclamatives" in terms of Zanuttini & Portner (2003). The result of this paper suggests that the presupposition of exclamatives can be captured at the syntactic level, endorsing the spirit of Zanuttini & Portner (2003).

Acknowledgement

Katsumasa Ito
University of Tokyo /
Japan Society for the Promotion of Science
k_ito@phiz.c.u-tokyo.ac.jp

References

Abels, K. (2010). Factivity in exclamatives is a presupposition. *Studia Linguistica* 64, pp. 141–157.

Antomo, M. & M. Steinbach (2010). Desintegration und Interpretation: Weil-V2-Sätze an der Schnittstelle zwischen Syntax, Semantik und Pragmatik. *Zeitschrift für Sprachwissenschaft* 29, pp. 1–37.

Blümel, A. (2012). Towards a novel analysis of was-für-split. *Proceedings of ConSOLE XIX*, pp. 105–123.

Bošković, Ž. (2005). On the locality of left branch extraction and the structure of NP. *Studia linguistica* 59:1, 1–45.

Chernilovskaya, A., Condoravdi, C., & Lauer, S. (2012). On the discourse effects of wh-exclamatives. *Proceedings of the 30th West Coast Conference on Formal Lingsuistics*, pp. 109–119.

Chomsky, N. (2000). Minimalist inquiries. Martin R., D. Michaels & J. Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, MIT Press, Cambridge (Mass.), pp. 89–156.

Chomsky, N. (2008): On phases. Freidin, R., C. P. Otero & M. L. Zubizaretta (eds.), *Foundational Issues in Linguistic Theory*, MIT Press, Cambridge (Mass.), pp. 133–166.

Corver, N. (1990). *The Syntax of Left Branch Extractions.* Doctoral dissertation, Katholieke Universiteit Brabant.

d'Avis, F. (2013). Exklamativsatz. In: Meibauer, J., Steinbach, M. & Altmann, H. (eds.), *Satztypen des Deutschen*, de Gruyter, Berlin, pp. 170–202.

Driemel, I. (2015). Exclamatives and Factivity: A new test based on VERUM Focus. *Proceedings of ConSOLE XXIII*, pp. 406–426.

von Fintel, K. (2004). Would you believe it? The king of France is back! Presuppositions and truth-value intuitions. Reimer, M. & A. Bezuidenhout (eds.), *Descriptions and Beyond*, Oxford University Press, Oxford, pp. 269–296.

Grimshaw, J. (1979). Complete Selection and the Lexicon. *Linguistic Inquiry* 10, pp. 279–326.

Haegeman, L. (2014). Locality and the distribution of main clause phenomena. Aboh, E. O., M. T. Guasti & Roberts, I. (eds.), *Locality*, Oxford University Press, Oxford, pp. 186–222.

Höhle, T. (1992). Über Verum-Fokus im Deutschen. Jacobs, J. (ed.), *Informationsstruktur und Grammatik*, Westdeutscher Verlag, Wiesbaden, pp. 112–141.

Lohnstein, H. (2000). Satzmodus—kompositionell. Akademie Verlag, Berlin.

Meinunger, A. (2004). Verb position, verbal mood and the anchoring (potential) of sentences. Lohnstein, H. & S. Trissler (eds.), *The Syntax and Semantics of the Left Periphery*, de Gruyter, Berlin, pp. 313–341.

Miyagawa, S. (2011). Genitive subjects in Altaic and specification of phase. *Lingua* 121(7), pp. 1265–1282.

Miyagawa, S. (2012). Agreements that occur mainly in the main clause. Aelbrecht, L., L. Haegeman & R. Nye (eds.), *Main Clause Phenomena: New Horizons*, Benjamins, Amsterdam, pp. 79–111.

Morzycki, M. (2012). The several faces of adnominal degree modification. *Proceedings of the West Coast Conference on Formal Linguistics* 29, pp. 187–195.

Reis, M. (2013). „*Weil*-V2"-Sätze und (k)ein Ende? Anmerkungen zur Analyse von Antomo & Steinbach (2010). *Zeitschrift für Sprachwissenschaft* 32, pp. 221–262.

Rett, J. (2008). A degree account of exclamatives. *Proceedings of SALT* 18, pp. 601–618.

Rett, J. (2011a). Exclamatives, degrees and speech acts. *Linguistics and Philosophy* 34:5, pp. 411–442.

Rett, J. (2011b). Exclamatives, degrees and speech acts. [Ms]. UCLA.
http://semanticsarchive.net/Archive/zUyMTY2Z/rett%20exclamatives%20LP%202011.pdf

Romero, M. & C. Han (2004). On negative yes/no questions. *Linguistics and Philosophy* 27:5, pp. 609–658.

Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics* 1, pp. 75–116.

Rosengren, I. (1992). Zur Grammatik und Pragmatik der Exklamation. Rosengren, I. (ed.), *Satz und Illokution Bd. 1*, Niemeyer, Tübingen, pp. 263–306.

Rossou, A. (2010). Selecting complementizers. *Lingua* 120:3, pp. 582–603

Solt, S. (2014). Q-adjectives and the semantics of quantity. *Journal of Semantics* 32:2, 221-273.

Truckenbrodt, H. (2006). On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics* 32, pp. 257–306.

Watanabe, A. (1993). Larsonian CP-recursion, factive complements, and selection. *Proceedings of the North East Linguistic Society* 23, pp. 523–537.

Zanuttini, R. & P. Portner (2003). Exclamative clauses: At the syntax-semantics interface. *Language* 79:1, pp. 39–81.

Deriving the Past Tense augment

Natalia Pavlou

Verbal morphology in the Greek past tense morphology involves the prefix *e-*, traditionally known as past tense augment, which is often taken to depend on stress placement and the prosody-morphology interface. As recently argued in Spyropoulos & Revithiadou (2009), this augment is a segmentally empty prefix with lexically encoded stress, realized after fission or when another formative does not satisfy the $T_{[+PAST]}$ node. This paper presents new data from Cypriot Greek simple verb forms and verbal complexes showing that the existing view of the augment as a stress exponent does not apply in this variety of Greek. The data from this non-standard variety of Greek contribute to a better understanding of the past tense augment by re-defining a morphological analysis that explains the phenomenon in Cypriot Greek.

1. Introduction

Tense formation in Standard Modern Greek, as well as Cypriot Greek[1], shows a two-way distinction, identified as past and nonpast. As reported in Holton et al. (1997), the future is formed periphrastically in Greek, *i.e.* it consists of a two-word phrase rather than a single word as in other languages. This periphrasis involves the element *θa* 'will' accompanied by the verb in the nonpast in Standard Modern Greek and *en na* in Cypriot Greek (Merchant & Pavlou 2017). Setting aside any interesting properties of the periphrastic formation of tense, this paper focuses on the morphology of tense in Greek: The distinction marked morphologically on the verb indicates the past and nonpast tense independently of other elements that combine with the verb syntactically.

---

[1]The data reported here are based on judgments from the author, who is a native speaker of Cypriot Greek. Any regional or other dialectal variation between speakers is not included here.

(1) *Non-past*

    a. ðe -n   -o
        tie -IMPERF -NONPAST-1SG
        'I am tying'

    b. ðe -s   -o
        tie -PERF -NONPAST-1SG
        'I will tie'

As seen above, the verb marks aspect with alternations in perfective and imperfective and tense and person agreement surface as a single morpheme for the 1ST PERSON SINGULAR. Past tense is morphologically marked on the verb's morphology, and differs from (1) by marking tense with different suffixes and the augment *e-*, as shown in the examples below.

(2) *Past*

    a. e-   ðe -n   -a
        PAST- tie -IMPERF -PAST-1SG
        'I was tying'

    b. e-   ðe -s   -a
        PAST- tie -PERF -PAST-1SG
        'I tied'

These examples show that Past Tense in Standard Modern Greek is marked morphologically on the verb with different exponents based on person and the appearance of the augment *e-*. Comparing Standard Modern Greek and Cypriot Greek, the augment appears in the latter and not the first.

(3) a. ðjava -s   -a
        read -PERF -PAST-1SG
        'I read'   Standard Modern Greek

    b. e-   ðjava -s   -a
        PAST- read -PERF -PAST-1SG
        'I read'   Cypriot Greek

As will be discussed in Section 3, this contradicts the idea that the augment develops for the antepenultimate stress to be attained (van Oostendorp 2012) or that the antepenultimate stress is an exponent of the [+PAST] as a segmentally empty prefix with lexically encoded stress (Spyropoulos & Revithiadou 2009). The paper investigates precisely the morphology of these two components: the past tense suffixes and the augment to argue against previous analyses that characterize the morpho-syntactic representation of the Standard Modern Greek verb as possible explanations of the verbal morphology in Cypriot Greek.

    Section 2 of the paper first discusses past tense suffixation and proposes a Distributed Morphology approach to tense and agreement suffixes. This proposal does not allow the adaptation of previous proposals to derive past tense morphology and the augment *e-* in Standard Modern Greek with *fission* (Spyropoulos & Revithiadou 2009).

    Second, Section 3 of the paper provides the data on the Cypriot Greek augment and proposes the underlying morphosyntactic structure involved in the Cypriot Greek verb. The Cypriot Greek augment does not have lexically encoded properties, an observation supported by the distribution of the augment in verbal complexes.

2. Past Tense suffixation

To provide a more general picture of the distribution of the Past Tense morphemes in this variety of Greek, this section presents the forms of the verbs in different verb classes. In Table 1, *psín-* 'to cook' belongs in the first conjugation and *fil-* 'to kiss' belongs in the second conjugation (first class). I provide the full paradigms below that show the different agreement and past tense suffixes.

Cypriot Greek Past Tense- Active Voice				
Person	PAST, PERF	Meaning	PAST, PERF	Meaning
1sg	é-psi-s-a	'I cooked'	e-fíli-s-a	'I kissed'
2sg	é-psi-s-e-s	'you cooked'	e-fíli-s-e-s	'you kissed'
3sg	é-psi-s-en	'he/she cooked'	e-fíli-s-en	'he/she kissed'
1pl	e-psí-s-a-men	'we cooked'	e-filí-s-a-men	'we kissed'
2pl	e-psí-s-e-te	'you cooked'	e-filí-s-e-te	'you kissed'
3pl	e-psí-s-a-n	'they cooked'	e-fíli-s-a-n	'they kissed'
3pl	e-psí-s-a-sin	'they cooked'	e-filí-s-a-sin	'they kissed'
Person	PAST, IMPERF	Meaning	PAST, IMPERF	Meaning
1sg	é-psi-n-a	'I was cooking'	e-fil-ú-s-a	'I was kissing'
2sg	é-psi-n-e-s	'you were cooking'	e-fil-ú-s-e-s	'you were kissing'
3sg	é-psi-n-en	'he/she was cooking'	e-fil-ú-s-en	'he/she was kissing'
1pl	e-psí-n-a-men	'we were cooking'	e-fil-ú-s-a-men	'we were kissing'
2pl	e-psí-n-e-te	'you were cooking'	e-fil-ú-s-e-te	'you were kissing'
3pl	e-psí-n-a-n	'they were cooking'	e-fil-ú-s-a-n	'they were kissing'
3pl	e-psí-n-a-sin	'they were cooking'	e-fil-ú-s-a-sin	'they were kissing'

The tables above present the distribution of the past tense suffixes in the active and nonactive voice. The verbal paradigm shows consistent appearance of the augment *e-* in both two-syllable and three-syllable verbs or in other words, those that have an already existing antepenultimate syllable or not. As will be discussed in section 3, this generalization contradicts the common assumption that the augment develops for the antepenultimate stress to be attained or that the antepenultimate stress is an exponent of the [+PAST] as a segmentally empty prefix with lexically encoded stress.

Before discussing the augment, I will discuss here the suffixes on the verb that also appear to be exponents of [+PAST]. The question that arises is whether these suffixes are portmanteau suffixes for both tense and person, as also argued in Joseph & Smirniotopoulos (1993) for mediopassive morphemes in Standard Modern Greek, or whether they can be exponents of different nodes, as already presented in the tables above. For example, the first assumption would take *-es* to be the realization of tense and person, while the second approach would take *-e* as the realization of tense and *-s* as the realization of person agreement. The latter is built on the assumption that an underlying syntactic structure provides a one-to-one matching of the morphemes with morphosyntactic features.

In the derivational framework of Distributed Morphology presented here, terminal nodes are projections of morphosyntactic features in an underlying syntactic structure and are realized as phonological exponents after Vocabulary Insertion (Arregi & Nevins 2012; Embick & Noyer

Cypriot Greek Past Tense- NonActive Voice		
Person	PAST, PERF	Meaning
1sg	e-plí-θ-ik-a	'I was washed'
2sg	e-plí-θ-ik-e-s	'you were washed'
3sg	e-plí-θ-ik-en	'he/she was washed'
1pl	e-pli-θ-ík-a-men	'we were washed'
2pl	e-pli-θ-ík-e-te	'you were washed'
3pl	e-plí-θ-ik-a-n	'they were washed '
3pl	e-pli-θ-ík-a-sin	'they were washed '

Person	PAST, IMPERF	Meaning
1sg	e-plin-ísk-u-mun	'I was being washed'
2sg	e-plin-ísk-e-sun	'you were being washed'
3sg	e-plin-ísk-e-tun	's/he was being washed'
1pl	e-plin-isk-ú-mastan	'we were being washed'
2pl	e-plin-ísk-e-stun	'you were being washed'
3pl	e-plin-ísk-u-ndan	'they were being washed'

Table 1: Cypriot Greek verbal morphology in NonActive

2007; Halle & Marantz 1993). I assume that morphology works on a set of terminal nodes which represent the relevant grammatical features to be encoded. The terminal nodes are organized in a hierarchical structure, which derives from the organization of the functional categories in the syntax and it is the result of the mapping of the syntactic structure onto the exponents with Vocabulary Insertion rules after Spell-Out. In this sense, I will pursue here an analysis where tense and agreement form separate terminal nodes.

A careful look of the surface forms in Table 1 and Table 2 show a consistent pattern: the suffix -*e* appears in 2ND person and 3RD SING. Other suffixes are -*a* and -*u* with both appearing in 1ST PERSON and 3RD PL. Other combinations create unattested forms of the verb:

(4) a. é-   psi   -s   -e   -s
         PAST- cook -PERF -PAST -2SG
         'You cooked'

     b. * é-   psi   -s   -a   -s
           PAST- cook -PERF -PAST -2SG
         (Int. 'You cooked')

     c. é-   psi   -s   -a   -men
         PAST- cook -PERF -PAST -1PL
         'We cooked'

     d. * é-   psi   -s   -e   -men
           PAST- cook -PERF -PAST -1PL
         (Int.'We cooked')

Given the systematic alternation between these exponents, the change can be the result of morphological operations that include feature-deletion rules yielding syntactic neutralization

in surface forms. In the active voice, *-e* is clearly not associated with any author features with the common property between the two exponents being the fact that they appear in Past Tense. I will assume here that person features are as follows.

(Noyer 1992; Halle 2000; Arregi & Nevins 2007)

(5)   a. [+Author] true iff the reference set contains the speaker
      b. [+Participant] true iff the reference set contains one of the discourse participants
      c. 1st: [+participant,+author]
      d. 2nd: [+participant,-author]
      e. 3rd: [-participant,-author]
      f. 3rd: [-participant,+author]=logically impossible

The following entries are then associated with the tense suffixes.

(6)   a. -a $\leftrightarrow$ [+past]
      b. -e $\leftrightarrow$ [-author +past]

Impoverishment (Bonet 1991), an operation that changes the featural content of morphemes prior Spell-out, deletes certain morphosyntactic features in certain contexts. To explain the distribution of the different exponents in the past, I propose that a deletion rule applies in 3RD PL and targets the deletion of the -author feature. As a result, *e-* cannot be inserted and *a-* is inserted as the less specified morpheme (e.g. *e-psi-s-a-n*, *e-psi-s-e-n 'they cooked').

(7)   [-author] $\rightarrow \emptyset$/___[+past]

The distribution of *u-* and *e-* in the nonactive imperfective will then work in a similar way, but only specified to happen in the environment specified below since *u-* would never appear in the active.

(8)   a. [+past] $\leftrightarrow$ -u /Voice$_{[-act]}$ Aspect$_{[-perf]}$___ [{+1ST, +3PL}]
      b. [-author +past] $\leftrightarrow$ -e

Impoverishment targets the deletion of the -author feature, resulting in the use of the less marked morpheme, namely *u-* in 3RD PL (e.g. *e-fiy-u-ndan*, *e-fiy-e-ndan 'they were kissing'). These assumptions are a way to explain the systematicity observed in the distribution of what is identified here as past tense suffixes.

    With this analysis, there is a possible way to account for the distribution of the Tense morphemes in the past without assuming that a single exponent realizes two morphemes, that is that T realizes both Tense and agreement. In previous work, Spyropoulos & Revithiadou (2009) have argued that the augment is derived via fission targeting the fused terminal node of T marking [α agreement, (+past)]. When the [+past] feature is not discharged, fision realizes this feature on another T position, surfacing as the augment. Provided this proposal, it follows that the assumption that fission targets the fused agreeement-tense node in the derivation of the past tense morphology cannot be implemented since the two nodes involved are no longer adjacent. Spyropoulos & Revithiadou (2009) also argue that the empty vocalic slot of the augment materializes under certain conditions providing a phonological form of this default exponent

of PAST. These conditions are built on the assumption that agreement and tense form a fused terminal node, which is subject to fission and that the empty prefix discharges this specification, which is the [+past] information. This does not clearly predict the restrictions of the prosodic structure, which only allow the augment *e-* to appear as a stressed antepenultimate syllable in Standard Modern Greek. Secondly, this analysis would not explain the facts, if agreement and tense were indeed separate projections, as argued here.

3. The Past Tense augment

3.1. Deriving the Cypriot Greek augment

The Past Tense augment appears only in some verbs in Standard Modern Greek, as conditioned by the number of syllables and the position of the stress (commonly called the antepenultimate stress (APU)). In this section, I show that this condition does not apply for the augment in Cypriot Greek and argue that a morphological approach can best capture the facts.

Stress in Standard Modern Greek is traditionally described as "dynamic stress" and it is acoustically manifested as longer duration or higher amplitude of the stressed syllable compared to the unstressed syllable(s) (Arvaniti 1999). Primary stress is always marked on one of the last three syllables, but its position is sometimes affected by specific morphemes (i.e. genitive suffix *u-*, past etc.). The Past Tense augment *e-* is found in two-syllable verbs that are stressed on the penult in the NonPast Tense (as in examples in Spyropoulos & Revithiadou 2009; Merchant 2015; van Oostendorp 2012; Ralli 2003). In these cases, the augment surfaces to hold the stress that retracts leftwards to the antepenultimate syllable in the past (9b). If the verb already has an antepenultimate syllable, then the stress shifts to it and the augment *e-* does not surface, as in (9d).

(9) *Standard Modern Greek*

    a. To psi -n   -o.
       it cook -IMPERF -NONPAST
       'I cook it.'

    b. To *(é)- psi -n   -a.
       it PAST- cook -IMPERF -PAST
       'I was cooking it.'

    c. To ðjavá -z   -o.
       it read -IMPERF -NONPAST
       'I read it.'

    d. To (*e-) ðjáva -z   -a.
       it PAST- read -IMPERF -PAST
       'I was reading it.'

In their analysis of the Standard Modern Greek augment, Spyropoulos & Revithiadou (2009) propose that:

> "the default exponent of PAST is a segmentally empty prefix with lexically-encoded accentual properties. Under certain conditions, the empty vocalic slot of the prefix materializes" (Spyropoulos & Revithiadou 2009:3).

Contrary to this proposal, the augment in Cypriot Greek does not actually carry lexical-encoded accentual properties. The Cypriot Greek augment *e-* appears as the antepenultimate syllable in two-syllable verbs (10b) and seems at a first glance to serve again stress-related purposes, as in Standard Modern Greek. However, it also appears in three-syllable verbs where the antepenultimate syllable is stressed, but the augment remains unstressed, as in (10d).

(10) *Cypriot Greek*

    a. psi -n   -o   to.
        cook -IMPERF -NONPAST it
        'I cook it.'

    b. *(é-) psi -n   -a   to.
        PAST- cook -IMPERF -PAST it
        'I was cooking it.'

    c. θcáva -z   -o   to.
        read   -IMPERF -NONPAST it
        'I read it.'

    d. *(e-) θcáva -z   -a   to.
        PAST- read   -IMPERF -PAST it
        'I was reading it.'

The difference observed here concerns our understanding of the augment as either a morpho-phonological phenomenon or an exponent of Past Tense that depends on a particular morphosyntactic structure. Similarly, the use of the augment in Ancient Greek was conceptualized in a different way than the corresponding Standard Modern Greek use. As Joseph & Janda (1988) mention "the augment, therefore, must be considered to be present in the underlying morphological structure of Ancient Greek past tense forms; furthermore, its occurrence there is not linked to any phonological feature(s)"(Joseph & Janda 1988: 198). This point emphasizes exactly that augments in languages that have them should not necessarily be seen as a phonological phenomenon provided the existence of the augment throughout the paradigm and not only under certain phonological conditions. Joseph & Janda (1988), for example, add that in a similar way the German *ge-*, as in (11), "has remained a morphological rule, albeit one with greater phonological conditioning"(Joseph & Janda 1988:201).

(11) a. Ich komme mit dem bus
         I   come   with the bus
         'I am coming with the bus.'

      b. Ich bin mit dem bus ge-   kommen
         I   be with the bus AUG- come
         'I have come with the bus.'

In this sense, this case of the Cypriot Greek augment is especially interesting in documenting and providing an analysis for its appearance based on the morphosyntactic structure of the verb.

    Cypriot Greek shows a different distribution of the augment compared to its use in Standard Modern Greek in that the Cypriot Greek augment is a Past Tense exponent which surfaces independently of stress. It depends on both the root and the suffixes, and not the root alone. This cannot be seen in Cypriot Greek given the augment's obligatory presence, but the restrictions

in Standard Modern Greek show the pattern in the following example. Example (12) shows a one-syllable root and example (13) a two-syllable root.

(12)   a. stel -o
           send -NONPAST-1SG
           'I am sending.'

       b. e-   stil -a
           PAST- send -PAST-1SG
           'I sent.'

       c. stil -a   -me
           send -PAST -1PL
           'We sent.'

(13)   a. xore -v   -o
           dance -IMPERF -NONPAST-1SG
           'I am dancing.'

       b. xore -ps   -a
           dance -PERF -PAST-1SG
           'I danced.'

       c. xore -ps   -a   -me
           dance -PERF -PAST -1PL
           'We danced.'

In (12b), the suffix that follows the root is one syllable, but in (12c) the suffixes that follow the root form two syllables. The augment appears only in (12b), but it does not when the suffixes that follow the root form more than one syllable. This clearly shows that the augment *e-* in Standard Modern Greek is conditioned by the count of syllables of the stem, and not the root, as argued in Spyropoulos & Revithiadou (2009). On the other hand, (13) shows a two-syllable root with (13b) having a one-syllable suffix and the suffixes in (13c) forming two syllables. As expected, the augment does not appear in any of these, since these examples already consists of three syllables. The following rules predict the distribution of the augment in Cypriot Greek:

(14)   a. T[+PAST] $\rightarrow$ e/___C
       b. T$\rightarrow\emptyset$

By the Elsewhere principle, which is based on the featural content or context of insertion being devoid of information and acting as the default, vowel-initial verbs will not show *e-*, as in the following example:

(15)   a. aɣap ó.
           love NONPAST-1SG
           'I love.'

       b. (*e)- aɣápi s-   a.
           PAST- love PERF- PAST-1SG
           'I loved.'

The following trees illustrate the appearance of the augment *e-* in Cypriot Greek.

(16) é-   psi   -s   -e   -s
    PAST- cook -PERF -PAST -2SG
    'You cooked (it).'

(17)   a.

    b.
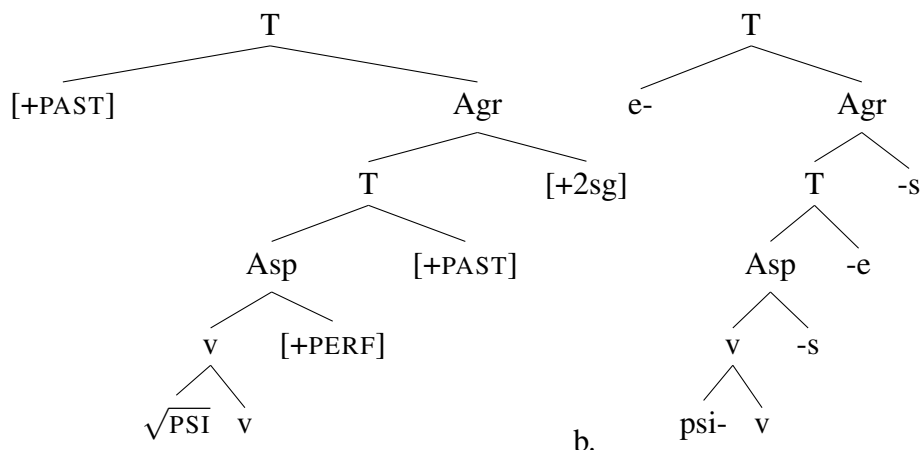
The Cypriot Greek augment *e-* is an exponent of [+PAST], as illustrated above, in an underlying structure where T and Agr form a fused terminal node.

3.2. *Tense and Aspect*

The differences in the syntactic distribution of the Past Tense augment between Standard Modern Greek and Cypriot Greek show that previous analyses that concern its co-occurrence with other morphemes in the structure also face problems with these data. It has been previously observed that the [+PERF] morpheme does not co-occur with the Past Tense augment in Standard Modern Greek independently of the number of syllables and the position of the stress.

(18)   a. stráf   -ik   -a
         throw-up -PERF -PAST
         'I threw up'

      b. * e-   stráf   -ik   -a
          PAST- throw-up -PERF -PAST
         (Int. 'I threw up')

      c. mb   -ík   -a
         enter -PERF -PAST
         'I entered.'

      d. * e-   mb -ík   -a
          PAST- enter -PERF -PAST
         (Int. 'I entered.')   Standard Modern Greek

The example in (18b) is not well-formed because of the Trisyllabic rule that allows stress to be positioned in a three-syllable window and *estrafika* consists of four syllables, assigning the stress on the antepenultimate syllable *straf-*. In (18d), however, the augment does not appear, even though it could form the antepenultimate syllable and hold the stress as predicted by the general rule found in the language. Spyropoulos & Revithiadou (2009) argue that the

	Root	Past, nonactive	Past,nonactive	Meaning
		Cypriot Greek perfective *-ik*, 2ND PERSON SINGULAR		
	skotó-	e- skotó -θ -**ik** -e -s	e- skotó-θ -∅ -i -s	'you were killed'
	fakk-	e- fatʃí -θ -**ik** -e -s	e- fatʃí-θ -∅ -i -s	'you were hit'
	sandano-	e- sandanó -θ -**ik** -e -s	e- sandanó -θ -∅ -i -s	'you were confused'
	aɣap	aɣap -í -θ -**ik** -e -s	aɣapí -θ -∅ -i -s	'you were loved'
h	*psi-*	e- psí -θ -**ik** -e -s	e- psí -θ -∅ -i -s	'you were cooked'
	vaf-	e- váf -t -**ik** -e -s	e- váf -t -∅ -i -s	'you were painted'
	pandrev-	e- pandréft -**ik** -e -s³	e- pandréft -∅ -i -s	'you were married'
	pe-	e- péxt -**ik** -e -s	e-péxt -∅ -i -s	'you shot yourself'
	mbe-	é- mb -**ik** -e -s	é-mb -∅ -i -s	'you entered'
	su-	e- súst -**ik** -e -s	e- súst -∅ -i -s	'you were shaken'
	ksev-	e- ksév -**ik** -e -s	e- ksév -∅ -i -s	'you got up on something'

Table 2: The perfective *-ik* in 2ND PERSON SINGULAR

competition between insertion of the augment and insertion of *-ik* and the win of the latter can be explained if it is seen as a more specified, in terms of features, morpheme, i.e. [+PAST, (+PERF)]. With both being past tense exponents, *ik-* is selected by perfective aspect. The augment is only specified for [+PAST], therefore insertion of *-ik* wins in perfective forms as the most specified morpheme.

In Cypriot Greek, the perfective *-ik* is optional[2] in the non-active forms of 2nd & 3rd SING. This clearly shows that *θik-* is not a single morpheme in Cypriot Greek (see Joseph & Smirniotopoulos 1993; Roussou 2009; Warburton 1973; Ralli 2005 for Standard Modern Greek) and that the augment has a morphological role in the underlying structure that is not affected by any other morpheme. This optionality seems to be the result of contextual allomorphy and the form of the tense morpheme. Realization of the tense exponent is sensitive to aspect, since the tense morpheme appears as *-e* or *-i*. In the cases where aspect surfaces null, the tense morphemes appear as *-i*; when aspect surfaces as *-ik*, then it surfaces as *-e* (see Pavlou (ms.)). Table 3 and Table 4 show a few examples, where the verb form optionally surfaces with *-ik* in 2ND and 3RD PERSON SINGULAR.

As expected, the Cypriot Greek augment distribution is not affected by the alternation from two- to three-syllable verbs stems.

(19) a. **e-**   stráf -**ik**   -e   -n
         PAST- return -PERF -PAST -AGR
         'He returned.'

    b. **e-**   stráf -i   -n
         PAST- return -PAST -AGR
         'He returned.'

---

[2]This is a marked option for Standard Modern Greek only for certain verbs that also show an unstressed past *e-* similar to Cypriot Greek, such as *e-θeáθ-i* 'was seen', *e-vréθ-i* 'was found', *e-ɣenníθ-i* 'was born', *e-kláp-i* 'was stolen', *i-kús-θ-i* 'was heard', *e-léx-θ-i* 'was said' etc. (Leivada, pc). In these examples, *-i* marks 3rd SINGULAR.

Cypriot Greek perfective *-ik*, 3RD PERSON SINGULAR *he/she/it*			
PAST, NONACTIVE	PAST, NONACTIVE		Meaning
skoto-	e- skotó-θ -**ik** -e -(n)	e- skotó-θ -∅ -i -(n)	's/he was killed'
fakk-	e- fatʃí -θ -**ik** -e -(n)	e- fatʃí-θ -∅ -i -(n)	's/he was hit'
sandano-	e- sandanó -θ -**ik** -e -(n)	e- sandanó -θ -∅ -i -(n)	's/he was confused'
aɣap-	aɣapí -θ -**ik** -e -(n)	aɣapí -θ -∅ -i -(n)	's/he was loved'
psi-	e- psí-θ -**ik** -e -(n)	e- psí-θ -∅ -i -(n)	'it was cooked'
vaf-	e- váf -t -**ik** -e -(n)	e- váf -t -∅ -i -(n)	's/he was painted'
pandrev-	e- pandré -ft -**ik** -e -(n)	e- pandré -ft -∅ -i -(n)	's/he was married'
pe-	e- péxt -**ik** -e -(n)	e- péxt -∅ -i -(n)	'it was played'
mbe-	é- mb -**ik** -e -(n)	é- mb -∅ -i -(n)	's/he entered'
su-	e- súst -**ik** -e -(n)	e- súst -∅ -i -(n)	'it was shaken'

Table 3: The perfective *-ik* in 3RD PERSON SINGULAR

    c. * stráf **ik**-   -e   -n
        return -PERF -PAST -AGR
        'He returned.'

    d. * stráf -i   -n
        return -PAST -AGR
        'He returned.'

The Vocabulary items for the insertion of *ik-* in Cypriot Greek given its distribution are as follows:

(20) a. Aspect[+PERF]↔ik/Voice[-ACT]___
      b. Aspect[+PERF]↔∅/Voice[-ACT]___-*i*

It is clear that *e-* does not compete with *-ik* in irregular verbs in Cypriot Greek, as in (21)[4].

(21) a. mbén -o
        enter NONPAS-1SG
        'I am entering.'
      b. <u>e-</u>   mb <u>-ik</u>   -a
        PAST- enter -PERF -PAST-1SG
        'I entered.'

This shows that an analysis where in the competition of the two, *-ik* is defined as the more specified morpheme and blocks insertion of the augment, is not supported and that *-ik* is in fact the exponent of perfective aspect.

---

[4]The stress in (21b) can surface either on the augment or the aspectual morpheme.

3.3.   The augment in verbal complexes

The augment *e-* can appear in two positions in verbal complexes[5] when the verb combines with an adverb[6]. In these cases, e_2- can be omitted, but when it is present and serves as the antepenultimate syllable, it holds the stress. The preverbs *psil-, mis-, poll-* are also free morphemes used as adjectives (e.g. *psilo xtirio* 'tall building') and *kal-, siɣ-* can be used as adverbs (e.g. *parpato siɣa* 'I walk slowly').

(22)   a. E_1-   psil-   o̱-   (e_2-)   psi   -s   -a   to.
             PAST- little- CM- PAST.1SG- cook -PERF -PAST.1SG it
             'I barely cooked it.'

        b. E_1-   mis-   o̱-   (e_2-)   psi   -s   -a   to.
             PAST- half- CM- PAST- cook -PERF -PAST.1SG it
             'I half cooked it.'

        c. E_1-   kal-   o̱-   (e_2-)   psi   -s   -a   to.
             PAST- good- CM- PAST- cook -PERF -PAST.1SG it
             'I cooked it well.'

        d. E_1-   poll-   o̱-   (e_2-)   psi   -s   -a   to.
             PAST- much- CM- PAST- cook -PERF -PAST.1SG it
             'I cooked it a lot.'

        e. E_1-   para̱- (e_2-)   psi   -s   -a   to.
             PAST- over- PAST- cook -PERF -PAST.1SG it
             'I overcooked it.'

        f. E_1-   siɣ-   o̱-   (e_2-)   psi   -s   -a   to.
             PAST- slow- CM- PAST- cook -PERF -PAST.1SG it
             'I slow-cooked it.'

In (22), the stress surfaces on the augment since the verb *psín-* 'to cook' belongs in the first conjugation where stress is found on the penult. In fact, when the stress is on the final syllable then stress shifts to the penultimate syllable as in *ksana-pon-ó > e_1-ksana-(e_2-)pón-u-n* 'I was in pain again'. So, stress always retracts a syllable to the left in the past. As seen in (23), the augment does not appear in non-past. The non-past expressed with *en na* is syntactically periphrastic in Cypriot Greek, characterized by the copula and a subordinating element (Merchant & Pavlou 2017, Terzi 1999) and uses a different set of suffixes compared to the Past Tense (i.e. *o-* for 1ST PERSON).

---

[5]When referring to the *e-* on the immediate left of the first member of the compound, it is indicated with e_1, while e_2 refers to the augment to the immediate left of the root. The underscore signifies the possible positions of the stress based on which one of the two positions acts as the antepenult.

[6]*-o-* is glossed as a *compound marker*, which is a linking vowel commonly found in compounds (Ralli & Karasimos 2009). The compound marker is inserted as *-o-* in most cases, independently of the gender of the noun that follows. For example, *kóptis* 'cutter' has masculine gender and *nix-o-kóptis* 'nail cutter' is marked with *-o-*. Similarly, a noun like *tsénda* 'bag', despite having feminine gender, also appears with *-o-* in a compound like *pay-ó-tsenda* 'old/useless bag'. When 'old' acts as an adjective to the noun, then the feminine suffix appears on both the adjective and the noun *paya tsenda* 'old bag'. Cases like these have been argued to be adverb incorporation cases in the VP (Rivero 1994). For English, see Bochnak 2013 for scalarity of 'half' in the VP.

(23) a. En na to psil- o- psí -s   -o.
          be C it little- CM- cook -PERF -NONPAST.1SG
          'I will barely cook it.'

      b. En na to mis- o- psí -s   -o.
          be C it half- CM- cook -PERF -NONPAST.1SG
          'I will half cook it.'

      c. En na to kal- o- psí -s   -o.
          be C it good- CM- cook -PERF -NONPAST.1SG
          'I will cook it well.'

      d. En na to poll- o- psí -s   -o.
          be C it much- CM- cook -PERF -NONPAST.1SG
          'I will cook it a lot.'

      e. En na to para- psí -s   -o.
          be C it over- cook -PERF -NONPAST.1SG
          'I will overcook it.'

      f. En na to siɣ- o- psí -s   -o.
          be C it slow- CM- cook -PERF -NONPAST.1SG
          'I will slow-cook it.'

There are no restrictions on the appearance of the augment in these cases in three syllable verbs, as is already predicted by the basic distribution seen in the previous section.

(24) a. E_1- psil- o- (e_2-) θkjáva -s   -a.
          PAST- little- CM- PAST- read   -PERF -PAST.1SG
          'I studied a little.'

      b. E_1- mis- o- (e_2-) θkjáva -s   -a.
          PAST- half- CM- PAST- read   -PERF -PAST.1SG
          'I studied enough, but not everything.' [lit. I half studied.]

      c. E_1- ksana- (e_2-) θkjáva -s   -a.
          PAST- again- PAST- read   -PERF -PAST.1SG
          'I studied again.'

Another case where the PAST can surface in two positions is with the use of *ksana* 'again', which also forms a compound with the verb. Unlike the cases seen above, the compound marker *o-* does not show up with the use of *ksana*, since the vowel *a-* is part of the root and a vowel does not need to be inserted between the two members of the compound. Unlike *ksana*, *a-* is not always part of the root and this can be seen by the appearance of the e_2-, as in *met-á-* 'after', e_1-*metá-lav-a* > e_1-*met-é$_2$-lav-a* 'to receive communion' and in *par-a-*, e_1-*par-á-lav-a* > e_1-*par-é$_2$-lav-a* 'to receive'.

(25) a. En na to ksana- psí -s   -o.
          be to it again- cook -PERF -NONPAST.1SG
          'I will cook it again.'

      b. E_1- ksan<u>a</u>- (<u>e_2</u>-) psi -s   -a   to.
          PAST- again- PAST- cook -PERF -PAST.1SG it
          'I cooked it again.'

NON-PAST, PERF	PAST,PERF	PAST,PERF, $+e_2$	Meaning
tsul-o-káts-o	e_1-tsul-ó-kats-a	e_1-tsul-o-$é_2$-kats-a	'to sit with knees bent'
ɣaur-o-mnjá-z-o	e_1-ɣaur-ó-mnja-s-a	e_1-ɣaur-o-$é_2$-mnja-s-a	'to look like a donkey'
anav-o-svín-o	anav-ó-svin-a	anav-o-$é_2$-svin-a	'to flicker'
tʃiy-o-pon-ó	e_1-tʃiy-o-pón-u-n	e_1-tʃiy-o-e_2-pón-u-n	'to have a stomachache'
xask-o-ɣel-ó	e_1-xask-o-ɣél-u-n	e_1-xasko-o-e_2-ɣél-u-n	'to gape and laugh'

Table 4: Verbal complexes

NON-PAST, PERF	PAST,PERF	PAST,PERF, $+e_2$	Meaning
xar-o-palé-fk-o	e_1-xar-o-pále-ps-a	e_1- xar-o-(e_2-)pále-ps-a	'to be at death's door'
ɣlik-o-kitá-z-o	e_1-ɣlik-o-kíta-ks-a	e_1- ɣlik-o-(e_2-)kíta-ks-a	'to have a sweet gaze'
strif-o-ɣirí-z-o	e_1-strif-o-ɣíri-z-a	e_1- strif-o-(e_2-)ɣíri-z-a	'to whirl around'
kli-o-stomnjá-z-o	e_1-kli-o-stómnja-s-a	e_1-klio-o-(e_2-)stómnja-s-a	'to stop talking'

Table 5: Verbal complexes with three-syllable verb stems

   c. En na to ksana- pé   -ps   -o.
      be to it again- send -PERF -NONPAST.1SG
      'I will send it again.'

   d. E_1-   ksana- (e_2-) pe   -ps   -a   to.
      PAST- again- PAST- send -PERF -PAST.1SG it
      'I sent it again.'

Double augments also appear in compounds[7] which show once again that the e_2- does not have to be stressed. In the following example, the stress is on the final syllable in the NonPast and retracts one syllable to the left in the Past. When e_2- is present, it is unstressed.

(26) a. I Patu tʃiy-   o- pon -á   kaθi mera.
       the Patu abdomen- CM- hurt -NONPAST every day
       'Patu has a stomachache every day.'

      b. I Patu e_1- tʃiy-   o- (e_2-) pón -e   -n   extes.
       the Patu PAST- abdomen- CM- PAST- hurt -PAST -AGR yesterday
       'Patu had a stomachache yesterday.'

The following tables show more examples with two-syllable & three-syllable verb stems, irregular and suppletive verbs that are part of N-V and V-V verbal complexes. The focus here is the distribution of the augment in two positions with nouns and verb serving as the first compound in the verbal complex.

---

[7]In Modern Greek, these V-V combinations are phonologically and morphologically true compounds since they have a single stress and a single inflection site, on the right edge of the second member. Further, the appearance of the linking vowel *-o* is exactly what is found in other compounds in Modern Greek (Nicholas & Joseph 2009).

The double appearance of *e-* confirms that it does not surface to only hold the stress in Cypriot Greek and that it is not affected by the classification of other prefixes. These data support the generalization for an unstressed augment in the past tense morphology of the verb, contrary to the idea that the augment bears lexically-encoded stress. The derivation of the reduplication facts of the augment remains a question for future work.

4. Conclusion

This squib has examined the distribution of the Past Tense augment in Cypriot Greek, a non-standard variety of Greek. The data presented here show that previous proposals on the derivation of the Past Tense augment in Standard Modern Greek do not capture the new facts presented here. Instead, a different approach is proposed to capture the distribution of the Cypriot Greek augment and the past tense suffixes.

Natalia Pavlou
University of Chicago
npavlou@uchicago.edu

References

Arregi, K. & A. Nevins (2012). *Morphotactics: Basque auxiliaries and the structure of spellout*, vol. 86. Springer Science & Business Media.

Arregi, K. & A. I. Nevins (2007). Obliteration vs. impoverishment in the Basque g-/z-constraint .

Arvaniti, A. (1999). Standard Modern Greek. *Journal of the International Phonetic Association* 29:02, pp. 167–172.

Bochnak, M. R. (2013). Two sources of scalarity within the verb phrase. *Studies in the composition and decomposition of event predicates*, Springer, pp. 99–123.

Bonet, E. (1991). *Morphology after Syntax: Pronominal Clitics in Romance*. Ph.D. thesis.

Dimela, E. & D. Melissaropoulou (2009). On prefix like adverbs in Modern Greek. *Patras Working Papers in Linguistics* 1, pp. pp–72.

Embick, D. & R. Noyer (2007). Distributed morphology and the syntax/morphology interface. *The Oxford handbook of linguistic interfaces* pp. 289–324.

Halle, M. (2000). Distributed morphology: Impoverishment and fission. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4* pp. 125–150.

Halle, M. & A. Marantz (1993). Distributed morphology and the pieces of inflection. the view from building 20, ed. by Kenneth Hale and Samuel Jay Keyser, 111–76.

Holton, D., P. Mackridge & I. Philippaki-Warburton (1997). *Greek: A comprehensive grammar of the Modern Language*. London: Routledge.

Joseph, B. D. & R. D. Janda (1988). The how and why of diachronic morphologization and demorphologization. *Theoretical morphology* pp. 193–210.

Joseph, B. D. & J. C. Smirniotopoulos (1993). The morphosyntax of the Modern Greek verb as morphology and not syntax. *Linguistic Inquiry* pp. 388–398.

Merchant, J. (2015). How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46:2, pp. 273–303.

Merchant, J. & N. Pavlou (2017). The morphosyntax of the periphrastic future under negation in Cypriot Greek. *Journal of Greek Linguistics* pp. 233–262.

Nicholas, N. & B. D. Joseph (2009). Verbal dvandvas in Modern Greek. *MIT working papers in linguistics: MIT workshop on Greek syntax and semantics*.

Noyer, R. R. (1992). *Features, positions and affixes in autonomous morphological structure*. Ph.D. thesis, Massachusetts Institute of Technology.

van Oostendorp, M. (2012). Stress as a proclitic in Modern Greek. *Lingua* 122:11, pp. 1165–1181.

Pavlou, N. (ms.). Allomorphic dependencies in verbal morphology. Ms., University of Chicago.

Ralli, A. (2003). Review article: Morphology in Greek linguistics: The state of the art. *Journal of Greek Linguistics* 4:1, pp. 77–129.

Ralli, A. (2004). Stem-based versus word-based morphological configurations: The case of Modern Greek preverbs. *Lingue e Linguaggio* 3:2, pp. 269–302.

Ralli, A. (2005). Morfologia [Morphology]. *Athens: Patakis* .

Ralli, A. (2010). Compounding versus derivation. *The Benjamins Handbook of Compounding. Philadelphia: John Benjamins Publishing Company* pp. 434–456.

Ralli, A. (2016). Affixoids: an intriguing category in morphology. 7th International Conference on Modern Greek Dialects and Linguistic Theory.

Ralli, A. & A. Karasimos (2009). The bare-stem constraint in Greek compound formation. *Gengo Kenkyu* 135, pp. 29–48.

Rivero, M. L. (1994). Clause structure and v-movement in the languages of the Balkans. *Natural Language & Linguistic Theory* 12:1, pp. 63–120.

Roussou, A. (2009). Voice morphology and ergativity in Modern Greek. *Proceedings of the 8th International Conference on Greek Linguistics (ICGL8)*, pp. 406–418.

Spyropoulos, V. & A. Revithiadou (2009). The morphology of past in Greek. *Studies in Greek linguistics* 29, pp. 108–122.

Terzi, A. (1999). Clitic combinations, their hosts and their ordering. *Natural Language & Linguistic Theory* 17:1, pp. 85–121.

Warburton, I. P. (1973). Modern Greek verb conjugation: Inflectional morphology in a transformational grammar. *Lingua* 32:3, pp. 193–226.

Clausal pied-piping in Basque wh-questions and syntactic optionality

Louise Raynaud

Basque embedded wh-questions exhibit apparent optionality between long-distance extraction of the wh-word and clausal pied-piping. This paper attempts to account for the pattern of free alternation found in Basque in a way that addresses the issue of syntactic optionality. It establishes a more comprehensive picture of the distribution of Basque clausal pied-piping in wh-questions, and shows that Cable's QP-based analysis of pied-piping can account independently for several restrictions, while providing the possibility of syntactic optionality. The central theoretical claim here is that a Q-based analysis is compatible with a Minimalist approach to optionality of the sort pursued in Biberauer & Richards (2006).

1. Introduction

Syntactic optionality is a bone of contention in the current Minimalist framework. It would seem rather intuitive that two different constructions that involve the same set of lexical items can be in free variation, i.e. be instances of semantically vacuous alternations. Yet this question takes on a particular weight in the Minimalist context. Indeed, the possibility of true optionality seems contrary to economy principles of the Minimalist Program, in which movement should always be motivated, either syntactically or at the interfaces.

And yet, such optionality is certainly found in natural languages. Biberauer (2003) describes, for example, the case of optional V2 in embedded sentences in Modern Afrikaans. Cable (2010b:167-170) provides many examples of optional DP-splits in several languages, among which German and Mohawk, along with instances of freely alternating preposition stranding and PP-pied-piping constructions in Icelandic, Irish and German. With regard to question strategies in particular, data has been provided for French, which famously displays a large range of question strategies. As highlighted in Adli (2006), Oiry (2011), Duguine & Irurtzun (2014) or Raynaud (2016) there are no semantic differences in presupposition requirements between wh-in-situ and wh-movement questions, which pattern alike. So it is a fact that some constructions alternate with no difference in interpretation.

While Basque has always been viewed as a bona fide wh-movement language, its wh-questions present some interesting peculiarities. Basque has the possibility to move a subordinate CP along with the wh-word in long-distance wh-questions. This operation, whereby a wh-word moves and drags along a larger constituent in which it is contained, in this case a whole clause, is named clausal pied-piping. Clausal pied-piping is a phenomenon that has only been observed in a handful of languages, such as Imbabura Quechua (Cole & Hermon 1981) or more recently Bangla (Simpson & Bhattacharya 2003). Pied-piping has been suggested to offer the 'clearest evidence of this indeterminacy in action' (Biberauer & Richards 2006:40), therefore the Basque constructions can be expected to provide insight into the question of optionality.[1]

This paper thus aims to show that true optionality in Basque wh-questions can be accounted for in a Minimalist framework. More specifically, I claim that a Q-based analysis such as Cable's provides the possibility of syntactic optionality, while accounting independently for several restrictions found in Basque. Based on previous work (Raynaud 2016), this paper also provide a more comprehensive picture of the distribution and the restrictions of clausal pied-piping in wh-questions. It does so through the lens of the question of syntactic optionality and shows that these constructions are instances of true optionality, i.e. that they are in free variation with their long-distance extraction counterparts. This can hopefully provide new insights in the controversial question of syntactic optionality in the Minimalist framework.

In section 2, I introduce the facts at hand about Basque wh-questions and argue that clausal pied-piping and wh-extraction are indeed in free variation in many contexts. Section 3 outlines cases in which this alternation is restricted by independent syntactic constraints such as islands. In section 4, it is argued that the optionality of these constructions can be accounted for by an analysis in terms of Q-particle (Cable 2010b). Such an analysis will be shown to be compatible with a Minimalist approach to optionality of the sort pursued in Biberauer & Richards (2006), who argue that while optionality is excluded from the functional motivation of movement, it is not excluded with respect to how a given feature can be formally satisfied. In other words, as long as formal requirements are fulfilled, 'the grammar doesn't mind' (Biberauer & Richards 2006). Section 5 demonstrates that a Q-based analysis accounts successfully for several independent restrictions as well. Finally, section 6 tries to provide an analysis that can explain the scope interactions and restrictions on pied-piping found with negation and interrogative complements.

2. Clausal pied-piping vs wh-word extraction: free alternation
2.1. Wh-question strategies in Basque

In Standard Basque, the usual strategy for forming questions is wh-movement. The questioned element, i.e. the wh-word, is required to move to the left edge of the clause. Movement of the wh-word triggers fronting of the verbal complex (V+Aux), resulting in a configuration where the wh-word is immediately left adjacent to the verb. The rest of the constituents follow the initial wh+verb group. From a neutral SOV word order (as in (1)), the order of constituents in a wh-question thus becomes: S_{WH}-V-O, as in (2), with the wh-word crucially sitting in a preverbal position.

---

[1]See also Duguine & Irurtzun (2014) for work on wh-in-situ in Labourdin Basque.

(1) Peiok   gereziak   jan ditu.
     Peio.ERG cherry.ABS.PL eat AUX
     'Peio ate the cherries.'

(2) **Nork**   jan ditu gereziak?
     who.ERG eat AUX cherry.ABS.PL
     'Who ate the cherries?'   (Duguine & Irurtzun 2014:1)

These constructions have generally been analyzed as involving movement of the wh-phrase to Spec,CP (represented by dotted lines), followed by movement of the verb complex to C, as an instance of T-to-C movement (represented by dashed lines) - in other words, a V2 effect (Ortiz de Urbina 1989; Arregi 2003; Heck 2008).

(3)   *Wh-movement and verb fronting*
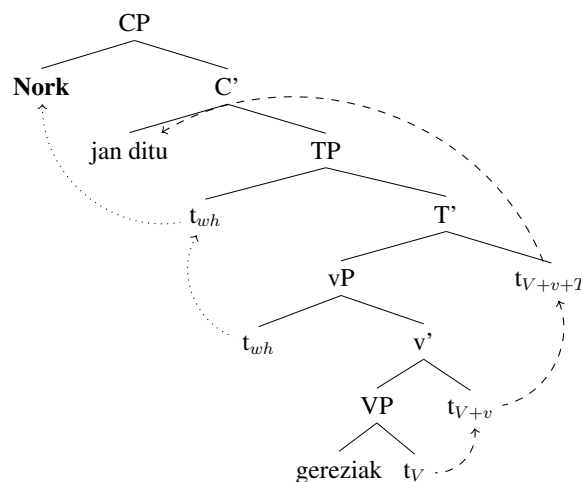
In these constructions, the position of the wh+verb group has to be clause-initial (Hualde & Ortiz de Urbina 2003), although topicalized elements may appear on the left field before the wh+verb group.

   The same strategy is applied in long-distance wh-questions which contain an embedded clause. It can then be observed that the wh-phrase moves through Spec,CP of the embedded clause to Spec,CP of the main clause. The cyclic nature of this movement through embedded Spec,CP can be evidenced by the fronting of the verbal complex of the embedded clause, as well as that of the main clause, as (4) illustrates, yielding a structure like (5).

(4) **Nork**   esan du Jonek [edan duela   ura]?
     who.ERG say AUX Jon.ERG [drink AUX.COMP water.ABS]
     'Who did Jon say drank water?'   (Duguine & Irurtzun 2014:2)

(5)   *Wh-movement in long-distance wh-questions*
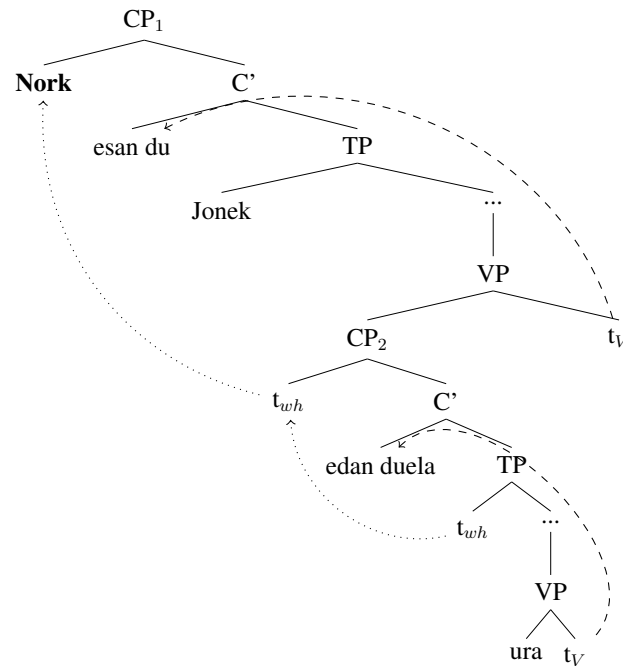
Interestingly, in Basque there is an alternative to long-distance extraction in wh-questions: clausal pied-piping. Pied-piping, an expression first introduced by Ross (1967), designates an operation by which an element moves and drags along a larger element in which it is contained. Small constituents like PPs or DPs may be pied-piped, but larger constituents can also be. This is the case in Basque, where in interrogative sentences, wh-words may pied-pipe an entire clause (CP) in which they are embedded. Consider the minimal pair formed by (4), repeated here for convenience as (6), and (7).

(6)   *Long-distance wh-extraction*

    Nork   esan du Jonek   [edan duela   ura]?
    who.ERG say   AUX Jon.ERG [drink AUX.COMP water.ABS]

    'Who did Jon say drank water?'

(7)   *Clausal pied-piping*

    [**Nork**   edan duela   ura]   esan du Jonek?
    who.ERG drink AUX.COMP water.ABS say   AUX Jon.ERG

    'Who did John say drank water?'   (Duguine & Irurtzun 2014:2-3)

In (7), the whole clause containing the wh-phrase appears in the same position as the wh-phrase would in long-distance wh-movement, i.e. to the immediate left of the matrix verb. In other words, it appears as if the wh-phrase has dragged along its whole clause. This process has been analyzed by Ortiz de Urbina (1989, 1993), Arregi (2003) and others as clausal pied-piping.

2.2. *Free alternation*

Basque embedded wh-questions exhibit optionality between long-distance extraction of the wh-word and clausal pied-piping. Several examples discussed in the literature show pairs in which both long-distance extraction (a. examples) and clausal pied-piping (b. examples) are possible.

(8) a. **Se** pentzate su [idatzi rabela   Jonek]?
          what think   AUX written AUX.COMP Jon.ERG

    b. [**Se** idatzi rabela Jonek] pentzate su?
       what written AUX Jon.ERG think   AUX
       'What do you think Jon wrote?'   (Arregi 2003:118)

(9) a. **Nor** esan du   Jonek   [joango dela]?
       who say AUX Jon.ERG go   AUX.COMP

    b. [**Nor** joango dela]   esan du   Jonek?
       who go   AUX.COMP say AUX Jon.ERG
       'Who has John said will go?'   (Ortiz de Urbina 1989:254)

(10) a. **Nor** nahi duzue (zuek)   [etor dadin]?
        who want AUX   (you.ERG.PL) come AUX.COMP

    b. [**Nor** etor dadin]   nahi duzue (zuek)?
       who come AUX.COMP want AUX (you.ERG.PL)
       'Who do you want to come?'   (Hualde & Ortiz de Urbina 2003:487)

(11) a. **Zenbat**   txakurrek   erran duzu [ausiki zaituztela]?
        how.many dog.ERG.PL say   AUX bite   AUX.COMP

    b. [**Zenbat** txakurrek   ausiki zaituztela] erran duzu?
       how.many dog.ERG.PL bite   AUX.COMP say   AUX
       'How many dogs did you say bit you?'   (online corpus Norantz)

These examples illustrate a variety of circumstances in which both the clausal pied-piping construction and the extraction version are acceptable. Clausal pied-piping is a 'quite pervasive' phenomenon, as stressed by Ortiz de Urbina (1989:249), that may occur with several types of subordination. The grammatical possibility of alternation appears guaranteed in many different lexical and syntactic contexts. It should be noted, however, that although in these environments the alternation is free, there appears to be a general effect of superiority of wh-movement over its pied-piping counterpart. The movement option is overall preferred by native speakers, i.e. judged more acceptable than its pied-piping counterpart (see Raynaud 2016). Although I have no definitive justification for this preference, it could be explained by economy considerations – nevertheless if we establish that it comes to an equal cost in terms of syntactic economy, we might have to admit that it could be more costly in terms of cognition/processing. This issue is left for future research.

Having established the possibility of free alternation, the question arises then whether this alternation is motivated on semantic grounds. In cases when the choice of pied-piping or extraction is not syntactically constrained, do they have the same meaning, or in other words does the choice of one or the other construction imply semantic differences?

2.3.   *The semantics of clausal pied-piping*

Arregi (2003) gives an account of the semantic properties of Basque clausal pied-piping struc-
tures. It appears that Basque clausal pied-piping is not semantically motivated: clausal pied-
piping constructions are semantically equivalent to their long-distance movement counterparts,
with regard to their scope and presupposition properties.

2.3.1.   *Matrix scope: amount wh-phrases (zenbat 'how many')*

Arregi (2003) demonstrates that clausal pied-piping results in the formation of a matrix question
in which the embedded argument takes scope over the whole structure. Simpson & Bhattacharya
(2003) highlight the same property of Bangla clausal pied-piping. Arregi's argument starts from
the observation that when amount wh-phrases like *zenbat* 'how many' undergo long-distance
movement over a scope bearing element, such as an intentional verb like *pentzaten* 'think', then
it results in an ambiguous interpretation. Consider the following:[2]

(12)   a.   *Long-distance movement*

[Semat   argaski]$_1$ desiriu rau Jonek   [$_{CP}$ erakusti lagunai   t$_1$]?
[how.many picture]$_1$ decided AUX Jon.ERG [$_{CP}$ to-show friends.DAT t$_1$]
'How many pictures did Jon decide to show his friends?'

b.   *Clausal pied-piping*

[$_{CP}$ Semat   argaski erakusti lagunai]   desiriu rau Jonek   t$_{CP}$?
[$_{CP}$ how.many picture to-show friends.DAT] decided AUX Jon.ERG t$_{CP}$
'How many pictures did Jon decide to show his friends?'   (Arregi 2003:128)

For both the long distance movement and the clausal pied-piping options, there are the same
two competing interpretations:

(13)   a.   decide > many
What is the number *n* such that Jon decided to show *n*-many pictures to his friends?

b.   many > decide
What is the number of pictures such that Jon decided to show those pictures to his
friends?

These two possible readings can be elicited by looking at possible answers to the questions
in a disambiguating context. Arregi provides the following scenario:

(14)   After his trip to New York, Jon has decided to show some of the pictures he took to his
friends. Since he does not want to bore them with too many pictures, he has decided
that he will show only forty of them. Furthermore, he has also decided that, among the
ones he will show, he will include twenty specific ones that are particularly beautiful.
(Arregi 2003:129)

In this scenario, both (12a) and (12b) can be answered by 'either forty, which would cor-
respond to the reading of the question where *many* takes scope under *decide* (13a), or twenty,
which would correspond to the reading in which *many* takes scope over *decide* (13b)' (Arregi

---

[2]All data from Arregi (2003) is from a dialect of Basque spoken in the western coastal town of Ondarroa.

2003:129). The fact there there is apparently no difference between the ambiguous interpretations of both structures would imply that they are semantically equivalent, and, as Arregi proposes, that they have the same LF structure.

2.3.2. *Presupposition*

Second, Arregi shows that the Basque long-distance wh-movement and its clausal pied-piping variant have no strong presuppositionality requirement. Consider examples (15) and (16), constructed after Arregi (2003).

(15) *Long-distance movement*

> **Nori** pentsatzen du Mirenek [$_{CP}$ eman diola Jonek musua]?
> who.DAT think AUX Miren.ERG give AUX Jon.ERG kiss.ABS

> 'To who does Miren think that John gave a kiss?'

(16) *Clausal pied-piping*

> [$_{CP}$ **Nori** eman diola Jonek musua] pentsatzen du Mirenek?
> who.DAT give AUX Jon.ERG kiss.ABS think AUX Miren.ERG

> 'To who does Miren think that John gave a kiss?' (Raynaud 2016:32)

Both sentences have the same level of presupposition: they both assume that there is someone that Miren thinks that Jon kissed. In neither of them does the matrix sentence inherit the stronger presupposition that Jon kissed someone. Indeed, the wh-word takes scope over the matrix clause and that results on a question on the matrix verb, and not only on the embedded clause. This can be checked by preceding the sentence by the denial that Jon actually kissed someone, as in (17), which does not make either sentence unfelicitous.

(17) Jonek ez dio inori musua eman, baina Mirenek pentsatzen du
Jon.ERG not AUX anyone.DAT kiss.ABS give but Miren.ERG think AUX
norbaiti musua eman diola.
someone.DAT kiss.ABS give AUX.COMP
'Jon did not kiss anybody, but Miren thinks he kissed somebody.'

(Raynaud (2016:32) after Arregi (2003:128))

It can thus be concluded that clausal pied-piping and long-distance extraction do not exhibit semantic differences when it comes to presupposition or scope.

3. *Restricted alternations*

Although constructions involving clausal pied-piping seem to be able to optionally alternate with long-movement structures in a variety of syntactic and semantic contexts, it appears that clausal pied-piping is not completely free either – it obeys a certain number of independent constraints with respect to which elements can pied-pipe and be pied-piped.

3.1. Obligatory fronting of the wh-word

First of all, Basque presents the distinctive property that the subordinate wh-clause can only be pied-piped if the wh-word is fronted to the left periphery of that clause (Cable 2010b).

(18)   a. [$_{CP}$ **Nor**$_1$ joango dela   t$_1$] esan du   Jonek?
                who go   AUX.COMP   said AUX John
             'Who did John say will go?'

       b. *[$_{CP}$ Joango dela   **nor**] esan du   Jonek?
                go   AUX.COMP who said AUX John
             'Who did John say will go?'   (Cable 2010b:154)

In the clausal pied-piping examples in (18) above, Heck (2008:106-107) observes that in (18a) both the verb and the auxiliary appear to the left of the complementizer *-(e)la* (indicating movement of the verb from T to C). The wh-phrase *nor* appears to the left of this verbal complex. As described in section 2, movement of the embedded verb and left adjacency of the wh-word indicates that the wh-word must have transited via the specifier position of the embedded CP. In (18b) by contrast, the wh-word appears to the right of the verbal complex, yielding an ungrammatical question. So (18) shows that Basque doesn't seem to allow wh-clauses to be pied-piped by wh-words that are internal to the subordinate CP, i.e. that have not moved to a preverbal position.

What matters for grammaticality purposes is the left-adjacency of the wh-word with respect to the verbal complex, and the requirement that nothing intervenes between them. The fact that it is at the left edge of the clause does not seem to matter as much, as evidenced by the fact that sentences with topicalized elements on the left field before the wh+verb group are acceptable. It is the intervention of an element between the wh-word and the verb that is considered ungrammatical. This is especially visible with direct object wh-words like *nor* 'who'. Consider the case of a simple question like (19).

(19)   a. Auzokideek   atzo   **nor**   ikusi zuten?
             neighbor.ERG.PL yesterday who.ABS see   AUX

       b. **Nor**   ikusi zuten auzokideek   atzo?
             who.ABS see   AUX neighbor.ERG.PL yesterday
             'Who did the neighbors see yesterday?'

In (19b) the wh+verb complex is fronted at the very left of the clause whereas in (19a) it is preceded by other elements. However, both sentences display immediate adjacency between the wh-word and the verb and are equally well-formed.

3.2. Islands

Furthermore, there are cases in which extraction of the wh-word is not allowed: wh-words cannot move out of certain islands (Ross 1967). As will be shown in this section, pied-piping can be a constrained alternative in such cases, but it need not always be, and I will claim that pied-piping is not solely a default option.

Wh-words originating in a referential adjunct clause cannot be extracted and yield ungrammatical sentences. With a temporal adjunct clause like (20a), extraction is indeed ungrammatical. Its clausal pied-piping counterpart (20b) is preferred.

(20) a. ***Zer** izutu   zen erregea [entzun zenuenean]?
          what frightened AUX king   hear   AUX.COMP.when

  b. %[**Zer** entzun zenuenean]   izutu   zen erregea?
      what   hear   AUX.COMP.when frightened AUX king
      'What did the king become frightened when he heard?'

In (20a), the wh-word *zer* 'what' is extracted out of the embedded adjunct clause and moved to the front of the matrix clause. As is expected, extraction of the wh-word out of an island is illicit. This contrasts with (20b), in which the entire embedded adjunct moves to Spec,CP leaving *zer* within the island and thus circumventing the extraction ban by clausal pied-piping.

Other types of islands yield similar effects. For instance, wh-words cannot be extracted from a relative clause, as exemplified in (21a) below. In that case, like above, the pied-piping option is acceptable, as in (21b).

(21) a. ***Nork** irakurri du Peruk   [idatzi zuen   liburua]?
          who   read   AUX Peru.ERG write AUX.COMP book
          'Who did Peter read the book that wrote?'

  b. [**Nork** idatzi zuen   liburua] irakurri du   Peruk?
      who   write AUX.COMP book   read   AUX Peru.ERG
      'The book that who wrote has Peter read?'   (Ortiz de Urbina 1989:249, 252)

Basque wh-words originating in a transitive subject clause (22) or a complex DP (23) may not be extracted either.

(22) ***Non** esaten   du asko zutaz   [onartua izateak]?
      where say.IMPF AUX a.lot you.about admitted being?
      'Where does it say a lot about you being admitted?'
      lit. 'To be admitted where says a lot about you?' (Hualde & Ortiz de Urbina 2003:485)

(23) ***Se** ikusi su   [idatzi raben gixona]?
      what seen AUX written AUX man.ABS
      'What did you see the man who wrote?'   (Arregi 2003:117)

The survey conducted in Raynaud (2016) suggests that, unlike what might have been thought, in these two last configurations, clausal pied-piping is not felicitous either and does not offer speakers a way to get round the ineffability scenario that arises from the extraction problem in wh-questions islands (Richards 1997). Pied-piping does not act, in these cases, as a last resort solution to a blocked movement alternative. It does, however, in the case of blocked wh-movement out of referential adjunct clauses and relative clauses, as illustrated above in (20) and (21). In these cases, since extraction of the wh-word and thus long wh-movement are impossible due to independent syntactic restrictions, clausal pied-piping appears to be a constrained alternative. As stated by Hualde & Ortiz de Urbina (2003:489), clausal pied-piping is the 'only strategy available for constituents inside adjuncts [...] which may not be extracted'. In other words, clausal pied-piping is then obligatory and there seems to be no optionality at play here. In a way, this is also true of the two cases in (22) and (23), where none of the two options is a satisfying alternative and there is no functional alternation. Nevertheless, as we have seen from the free variation contexts, it is also not the case that clausal pied-piping is a 'default option' (Horvath 2006) that occurs if and only if the pied-piper cannot be extracted from the clause. The possibility for syntactic optionality therefore needs to be formally acounted for.

4.   Accounting for syntactic optionality

4.1.   The theoretical puzzle

The concept of true optionality is a controversial one. The criterion for optionality is surely 'phenomenological' in nature: if speakers have the choice between different word orders, then two structures stand in optionality with respect to one another. But although true optionality could be defined as 'semantically vacuous alternations in surface order' (Biberauer & Richards 2006:35), its syntactic and conceptual implications remain under discussion, because it poses a theoretical problem in the context of the Minimalist Program.

True optionality, of the sort that is observed in the above examples, appears to be contrary to economy principles of the Minimalist Program, namely Last Resort (LR) and Full Interpretation (FI), which regulate operations, ensuring that neither too much nor too little occurs. So concretely, nothing happens for no reason: everything that happens in a derivation must be motivated either by a formal requirement, such as feature checking, or by an interface requirement, i.e. a difference in meaning. This is the so-called Fox-Reinhart intuition on optionality, captured by Chomsky (2001:34) as a 'general economy principle': 'An optional rule can apply only when necessary to yield a new outcome'. In this spirit, Move/internal Merge is something that does not occur freely: to happen, it needs to be triggered by features. In the case of wh-movement, it has been argued that a wh- or Q-feature is borne by the wh-word and needs to be checked/valued against C, which carries the same feature. What follows from the Fox-Reinhart intuition is that given a possible alternation between two different constructions, one of them must be 'marked', in the sense that it is less economical. If they have the same meaning, then this alternation constitutes a violation of LR, as it implies non-feature-driven movement. What is more, it also constitutes a violation of FI if there is no interpretative effect ('α enters the numeration only if it has an effect on output'(Chomsky 1995:294)). Under these circumstances, there should not be a possible choice between moving and not moving, rendering the presence of semantically vacuous alternations superfluous, and thus impossible. 'Since movement is the operation that feeds the surface order of constituents, it follows that word-order alternations involving (what appears to be) the same set of lexical items should not exist.' (Biberauer & Richards 2006:36).

We can see how the facts that we just described about Basque clausal pied-piping constructions fit into this debate. Contrary to what is predicted by Minimalism, it seems that in some cases movement is indeed optional. How can the pattern of optionality found in Basque be accounted for in terms that are compatible with economy principles of Minimalism? Two sub-questions arise here. First of all, does the general architecture of Minimalism allows for such optionality conceptually and formally, and how? Section 4.2 aims at answering this question by showing that true optionality is actually built-in the Minimalist system, as demonstrated by Biberauer & Richards (2006), if we consider that such a system is driven by the satisfaction of feature checking requirements. These theoretical preliminaries established, the rest of this paper addresses the second question, i.e. how can we implement the Basque data concretely in the syntax while preserving the conceptual insights? I argue that Cable's Q-based analysis of pied-piping provides the possibility of syntactic optionality, while accounting independently for the alternation restrictions found in Basque. The central theoretical claim here is that a Q-based analysis is compatible with a Minimalist approach to optionality of the sort pursued in Biberauer & Richards (2006).

4.2. Built-in underdetermination: True Optionality

Among the propositions that have been made to account for optionality in a Minimalist framework, one possibility consists in saying that optionality – of the true kind – is actually built in the system.[3] The possibility of truly optional variation is theoretically legitimated by Biberauer & Richards (2006). While such true optionality would appear to be contrary to economy principles of Minimalism, Biberauer and Richards remark that such variation is actually admitted on system-internal grounds. They argue that whereas optionality is indeed excluded from the functional motivation of movement, it is not with respect to how a given feature is formally satisfied. In other words, sometimes 'the grammar doesn't mind' (Biberauer & Richards 2006) as long as formal requirements are fulfilled, and there arises true optionality.

When there are actually grounds for the system to privilege one option over the other, then it does, but if not, Biberauer and Richards argue, it can be more efficient to leave the choice open. Efficiency in the syntax is determined in terms of cost. Formerly, cost was determined on the basis of number of steps. In the more recent Agree-based framework however, it relies on a more local notion of feature-triggered operations. As a result, 'an operation Ω will now be equally as costly as any other operation Ω' that may potentially apply at a given stage σ of the derivation if Ω and Ω' are both valid ways of satisfying the formal imperative F driving operations at σ (i.e. both Ω and Ω' result in a well-formed structure, obey locality, etc.). LR and FI simply require that F be (immediately) satisfied; they do not specify how. Therefore, Ω and Ω' are optional operations with respect to each other.'(Biberauer & Richards 2006:39). As pointed out already above, as these operations are motivated by obligatory features that are already present in the derivation no matter what, they need not yield for a new interpretation, and can result in a semantically vacuous alternation. This approach introduces a definition of optionality that relies on the equal satisfaction of formal and featural requirements rather than on the identity of derivations.

4.3. A Q-based analysis

The Basque data can be successfully accounted for in the framework of Cable (2010b, 2012, 2013), which predicts this kind of free alternation, in a way that is consistent with Biberauer & Richards (2006)'s approach to optionality.

Cable proposes a novel account of pied-piping and wh-movement in general, by claiming that the wh-feature of the wh-word is not the actual target of wh-movement operations. Instead, he proposes the existence of a Q-particle, based on data from Tlingit, a Na-Dene language. Tlingit wh-questions require the wh-word (i) to precede the main predicate of the clause, like Basque, and (ii) to be followed by the interrogative particle *sá*, which either directly follows the wh-word or a phrase properly containing it (Cable 2010b, 2012).

(24) [[Wáa kwligeyi $_{CP}$] xáat $_{NP}$] **sá** i   tuwáa sigóo?
      how   it.is.big.REL   fish   Q do.you.want
      'How big a fish do you want?'
      lit. 'A fish that is how big do you want?'   (Cable 2010a:572)

---

[3]For other proposals, see for instance Adger & Smith (2005) for the idea of multiple grammars/parametrized variation and Müller (2001) for an account in terms of optional features.

(25) Daa **sá**wé i   éesh al'óon?
      what Q.FOC your father he.hunts.it
      'What is your father hunting?'

                                        (Dauenhauer & Dauenhauer (2000) qtd. in Cable (2010b:22))

Cable argues that this particle bears a Q-feature that must agree with C_Q, thereby triggering movement of the QP to Spec,CP. That Q-particle/operator is present in all languages and may be phonologically realized, as in Tlingit, or not, as in English or Basque. The Q-particle merges with the phrase containing the wh-word, taking this phrase as its complement and agreeing with it, as in (26). It is this Q-particle that is the target of wh-movement operations, instead of the wh-word itself.

(26)   *The Q-Based Analysis of Wh-Movement* (Cable 2012:823)
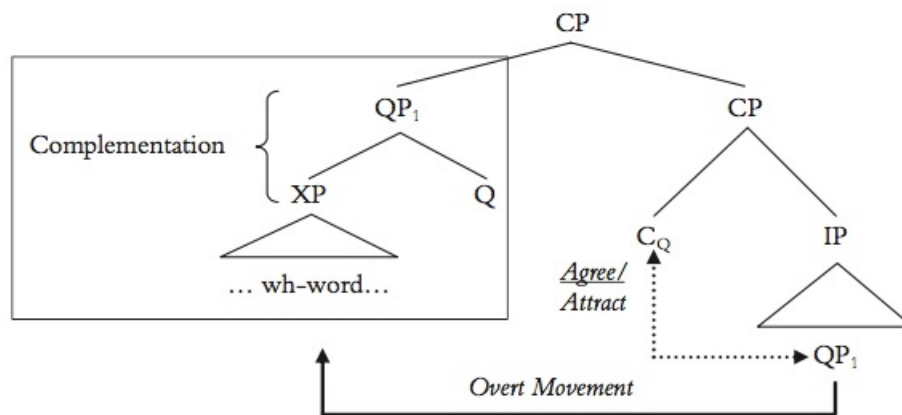
The existence of pied-piping structures is derived from cases where the targeted Q-particle is not directly structurally adjacent to the wh-word. Instead of merging with the DP directly containing the wh-word, the Q particle may merge with a bigger constituent containing it, such as a PP or a CP. In this case, the Q particle takes the whole embedded CP as its complement. The QP, containing the CP, is then attracted to the specifier position of the matrix CP to check its [+Q] feature against C_Q, thus moving the entire embedded clause. In this perspective, it is to be noted that this account actually negates the existence of pied-piping. Indeed, for Cable it is no longer the case that 'an operation that targets the features of a lexical item L applies to a phrase properly containing the maximal projection of L'(Cable 2012:817). Since the target of movement is no longer the wh-word, it does not drag along its maximal projection, and no pied-piping is involved. Cable talks of 'pied-piping structures', which he distinguishes from the operation of pied-piping itself. How can Cable's account be implemented in the case of Basque to show the possibility of true optionality? I will try to show that it naturally accounts for syntactic optionality in the sense established by Biberauer & Richards (2006), and how it deals with restrictions on the distribution of clausal pied-piping.

   In pied-piping structures, the Q-particle is not structurally adjacent to the wh-word. As illustrated in (26) above, in a structure involving extraction of the wh-word, the Q-particle is merged in a position directly adjacent to the wh-word and becomes its sister, taking it as its complement. In order for a clausal pied-piping structure to be possible, the QP must be merged in a

position that allows it to move the whole CP: therefore, the QP needs to take the CP as its complement. Indeed, theoretically nothing requires that 'Q be 'as close as possible' to the wh-word it associates with'(Cable 2010b:167). So nothing prevents the wh-word from being optionally extracted from its original clause or not. Moreover, in the same conditions where there are no barriers and no intervention effects, clausal pied-piping and wh-extraction can theoretically occur in free variation. Such an account correctly predicts the occurrence of alternations such as those found in Basque. In order to see how that works, recall examples (6) and (7) repeated here for convenience, which are in free variation, and observe their corresponding derivations.

(27) **Nork** esan du Jonek [edan duela ura]?
who.ERG say AUX Jon.ERG [drink AUX.COMP water.ABS]
'Who did Jon say drank water?'

(28) [**Nork** edan duela ura] esan du Jonek?
who.ERG drink AUX.COMP water.ABS say AUX Jon.ERG
'Who did John say drank water?' (Duguine & Irurtzun 2014:2-3)

(29) *QP-extraction in Basque*

$$
\begin{array}{l}
\text{CP}_1 \\
\quad \text{QP} \quad \text{C'} \\
\qquad \text{Nork} \quad \text{Q} \quad \text{C}_Q \quad \text{TP} \\
\qquad\qquad \text{esan du} \quad \text{Jonek} \quad \text{VP} \\
\qquad\qquad\qquad \text{CP}_2 \quad t_{V1} \\
\qquad\qquad\qquad t_{QP} \quad \text{C'} \\
\qquad\qquad\qquad\qquad \text{edan duela} \quad \text{TP} \\
\qquad\qquad\qquad\qquad\qquad t_{QP} \quad \text{vP} \\
\qquad\qquad\qquad\qquad\qquad\qquad t_{QP} \quad \text{VP} \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{ura} \quad t_{V2}
\end{array}
$$
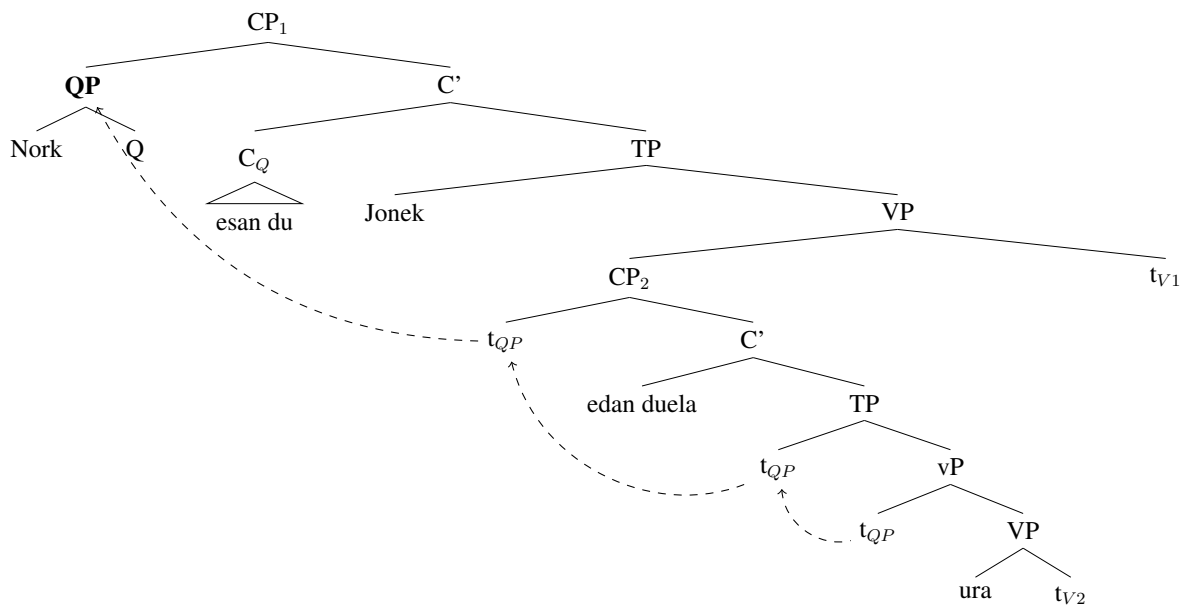
(30) *Q-based CP-pied-piping in Basque*

$$
\begin{array}{l}
\text{CP}_1 \\
\quad \text{QP} \quad \text{C'} \\
\qquad \text{CP}_2 \quad \text{Q} \quad \text{C}_Q \quad \text{TP} \\
\qquad \text{Nork} \quad \text{C'} \quad \text{esan du} \quad \text{Jonek} \quad \text{VP} \\
\qquad\qquad \text{edan duela} \quad \text{TP} \quad\quad t_{QP} \quad t_{V1} \\
\qquad\qquad\qquad t_{wh} \quad \text{vP} \\
\qquad\qquad\qquad\qquad t_{wh} \quad \text{VP} \\
\qquad\qquad\qquad\qquad\qquad \text{ura} \quad t_{V2}
\end{array}
$$
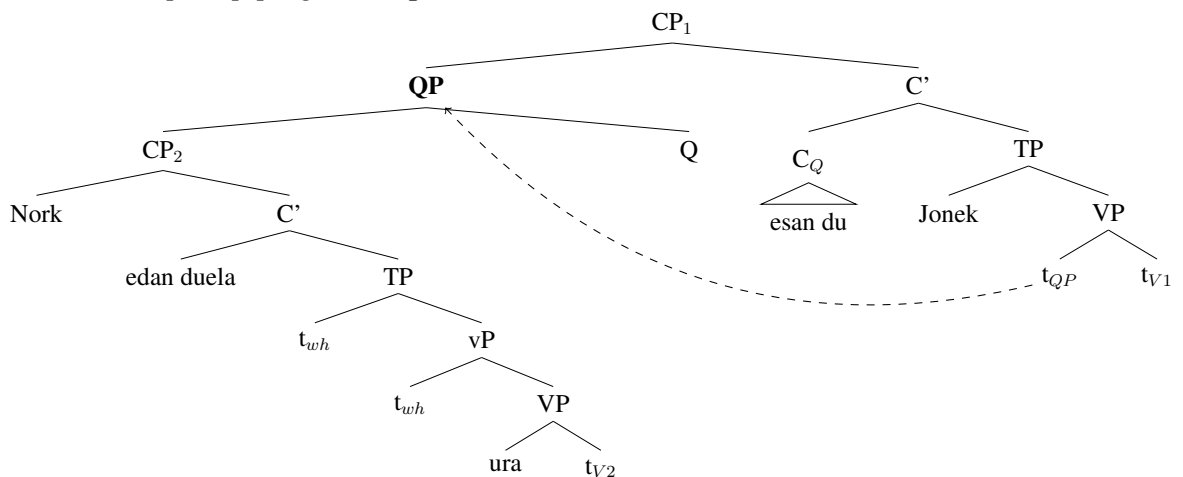
We can see how Biberauer and Richards's (2006) proposal is compatible with an analysis of pied-piping like the one offered by Cable. Derivations (29) and (30) both are adequate ways to satisfy their featural requirements (i.e. to check off the Q-feature). They have the exact same lexical array and differ only in where Q Merges. Therefore a Q-based account of wh-movement à la Cable provides us with a straightforward answer to the licensing of clausal pied-piping in a feature-based framework on the one hand and the possibility of syntactic true optionality on the other. Indeed, the insertion of Q in the derivation may apply at two different stages σ of the derivation and represent in both cases a valid way of 'satisfying the formal imperative F driving operations at σ'(Biberauer & Richards 2006:39). Both operations will result in a well-formed structure, illustrated by (27) and (28), which satisfy formal requirements and obey independent constraints such as locality. Therefore, Cable's analysis of clausal pied-piping provides us with a framework in which wh-extraction and clausal pied-piping can be said to be syntactically optional operations with respect to each other.

5.   *Accounting for restrictions*

Recall that it has been established that some contexts impose restrictions on the distribution of pied-piping and wh-extraction. There are some instances where one option is not available and the other prevails, like with adjuncts or relative islands that constitute independent syntactic constraints to movement. I will now show that Cable's Q-based analysis successfully accounts for these restrictions as well.

5.1.   *Q-Agreement*

In order to account for the locality constraints on pied-piping, Cable postulates that in some languages, pied-piping is limited by the need for the wh-word and the Q-particle to Agree. In these so-called Q/wh-Agreement languages, the Q-particle bears an unvalued Q-feature that must be valued through Agreement with a wh-word in its c-command domain.

(31)   *Q-wh Agreement* (Cable 2013:124)
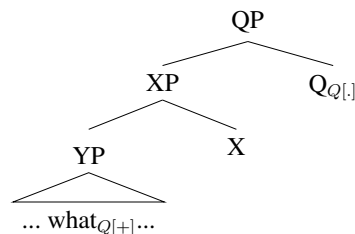
It then follows that if certain structures are barriers to Agreement, like islands, these will not be able to c-command a wh-word within the pied-piped phrase, and the resulting structure will be excluded. This leads to the following consequences. Cable (2010b) predicts that pied-piping past lexical categories is forbidden in Q/wh-Agreement languages, and more generally that pied-piping across separate phases should be impossible. Cable relies on the notion of derivational phase introduced by Chomsky (2001). Phases are lexical subarrays that get spelled-out in a

cyclic fashion, thus breaking up the derivation in smaller domains. Such a conception of the derivation necessarily affects the possibility of movement, as captured for instance by the Phase Impenetrability Condition (PIC).

(32) **The Phase Impenetrability Condition** (after Chomsky 2001)
The domain of [a phase head] H is not accessible to operations at [the next highest phase] ZP; only H and its edge are accessible to such operations.

Thus in virtue of the PIC, Cable states, the wh-word cannot be in a separate phase/Spell-Out domain from the head Q of the pied-piped phrase. Recall:

(33) a. [$_{CP}$ **Nor**$_1$ [$_{IP}$ joango dela t$_1$]]$_2$ esan du   Jonek t$_2$?
            who   go   AUX   said AUX John
      'Who did John say will go?'

      b. *[$_{CP}$ [$_{IP}$ Joango dela **nor**]]$_2$ esan du   Jonek t$_2$?
            go   AUX who   said AUX John
      'Who did John say will go?'   (Cable 2010b:154)

If we assume with Cable that Basque is a Q/wh-Agreement language, then we can account for a constraint we already mentioned on the internal form of pied-piped CPs in Basque: a clause can only be pied-piped if the wh-word has been moved to the left edge of that clause. It is impossible for a subordinate CP to be pied-piped by wh-words internal to it. Ungrammatical structures like (33b) follow from the inability for Agreement to apply across separate Spell-Out domains. Since C heads are a typical example of phase heads, a structure like (33b) would require agreement between Q and the wh-word to apply in separate Spell-Out domains, something that the PIC forbids. Thus, 'in order for a Q-particle to Agree with a wh-word contained within a subordinate CP, that wh-word must undergo movement to the specifier of the CP'(Cable 2010b:155). Indeed, Spec,CP is a phase edge and remains accessible for agreement after Spell-Out. Thanks to that movement, the wh-word and the Q-particle are placed in the same phase and can agree.

(34) *Blocked Q/wh-Agreement in Basque clausal pied-piping*
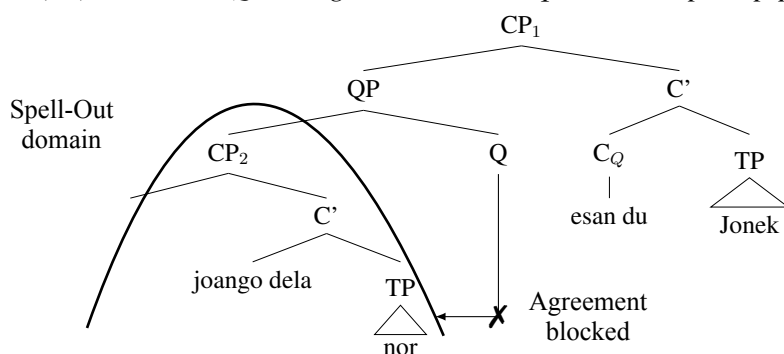
5.2.   QPs and islands

As we saw, there are further restrictions that apply to clausal pied-piping in Basque, and more specifically to its free alternation with wh-extraction. Firstly, there are cases in which movement of the wh-word out of the clause is ungrammatical, and clausal pied-piping is the only option.

These configurations (Subject Condition, Complex DPs, Adjuncts and Relative clauses) are islands, and wh-extraction is therefore expected to be impossible. What needs to be explained is how a Q-particle framework accounts for the possibility of clausal pied-piping as an alternative in these cases. Recall example (20):

(35) a. *__Zer__ izutu zen erregea [entzun zenuenean]?
     what frightened AUX king hear AUX.COMP.when

 b. %[__Zer__ entzun zenuenean] izutu zen erregea?
     what hear AUX.COMP.when frightened AUX king
     'What did the king become frightened when he heard?'

(35a) is ungrammatical because extraction of the wh-word out of the adjunct island is forbidden. (35b), its clausal pied-piping counterpart, is grammatical. In constructions that involve extraction of the wh-word, we know that the Q-particle is merged in a position directly adjacent to the wh-word which becomes a complement of the QP. Since islands are barriers for movement, it is ungrammatical for the QP to move and Agree with C_Q, as illustrated by (36)

(36) *Q-agreement blocked*
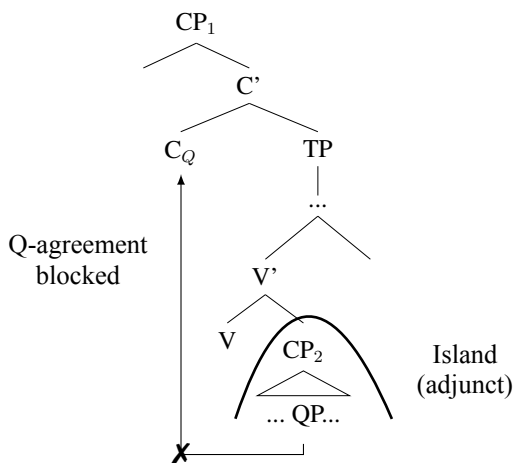
(37) *Q-agreement ok*
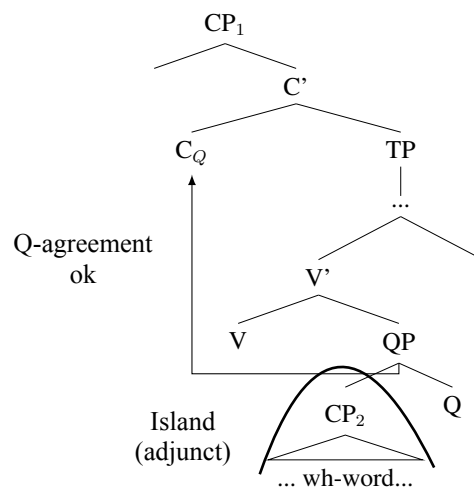
But in the clausal pied-piping configuration, the Q-particle merges in a position where it is no longer dominated by the phase boundary induced by the subordinate C, and movement of the QP no longer involves crossing a barrier and is therefore perfectly grammatical (37).[4]

So far, we have seen that Cable's account explains nicely cases in which the alternation is syntactically free. Indeed, as transpires from the present analysis of Basque clausal pied-piping constructions with respect to their long-distance movement counterpart, the numerations appear to be the same and the derivations appear to involve the same number of steps. Furthermore, Cable's Q-theory can also account for situations in clausal pied-piping is the constrained alternative when extraction of the wh-word is not permitted, and finally the necessity of the wh-word to be at the left edge of the clause. Finally, Cable's account has the advantage of predicting cases in which both wh-extraction and clausal pied-piping are viable options and can potentially alternate freely.

---

[4]This story, however, is not specific to Cable. Heck (2008) also resorts to the argument that wh-extraction can be blocked by an intervening island because of impossible Agreement. For him however, Agreement takes place between the wh-word and the interrogative C_{+WH}.

However, his account does not make any predictions about further cases of restricted alternation, in which clausal pied-piping is not allowed in Basque and to which I turn now.

6.　The scope problem

While the above accounts for the possibility of syntactic optionality, one might expect there to be semantic consequences associated with which element the Q-particle merges with. Although we have seen that there is no difference in interpretation (section 2.3), this other class of restrictions on the alternation of wh-extraction and clausal pied-piping would suggest some sort of a scope effect.

6.1.　Negation

In these cases, clausal pied-piping is ungrammatical, and extraction is therefore the only option. For instance, whereas long-distance extraction is allowed across negation (38a), clausal pied-piping is not (38b) (Arregi 2003).

(38)　a.　$Sein_1$ **es** tau esan Mirenek $[_{CP}$ t_1 jun danik]?
　　　　　Who not AUX said Miren.ERG $[_{CP}$ t_1 gone has]

　　　b.　*$[_{CP}$ Sein jun danik] **es** tau　　esan　　　　Mirenek t_{CP}?
　　　　　$[_{CP}$ who gone AUX] not AUX said Miren.ERG
　　　　　'Who didn't Miren say left?'　　　　　　　　　　　　　　(Arregi 2003:135)

Arregi (2003) accounts for this negative island effect by hypothesizing that clausal pied-piping involves reconstruction at LF and argues, after Heim (1992) and Beck (1996), that negation blocks reconstruction of pied-piped material. Concretely, he claims that the pied-piped complement reconstructs at LF in its base position, leaving the wh-word behind. This results in an LF structure which mirrors the long-distance extraction construction (as shown by the parallel structures in (39) and (40b)), thereby accounting for their semantic equivalence.

(39)　*LF structure for long-distance extraction*

　　　$[_{CP_1}$ $sein_i$ esan tau Mirenek $[_{CP_2}$ t_i jun danik]]
　　　$[_{CP_1}$ Wh_i ...　　　　　　　　$[_{CP_2}$ t_i ...　]]

(40)　a.　*Surface structure for clausal pied-piping*

　　　　　$[_{CP_1}$ $sein_i$ $[_{CP_2}$ t_i jun danik] esan tau Mirenek t_{CP2}]
　　　　　$[_{CP_1}$ Wh_i $[_{CP_2}$ t_i ...　　　] ...　　　　　　　t_{CP2}]

　　　b.　*LF structure for clausal pied-piping*

　　　　　$[_{CP_1}$ $sein_i$ t_{CP2} esan tau Mirenek $[_{CP_2}$ t_i jun danik]]
　　　　　$[_{CP_1}$ Wh_i t_{CP2} ...　　　　　　　$[_{CP_2}$ t_i ...　]]

This reconstruction happens in two steps: first, the wh-word moves out of the pied-piped clause, so it takes scope over the matrix sentence (40a). Second, the remaining clause is reconstructed to its base position, where it is interpreted, just as it would be in the long-distance counterpart (40b). The intervention of negation makes reconstruction impossible, and therefore excludes the possibility of clausal pied-piping in the presence of matrix negation (42).

(41)   *Long-distance extraction over negation*

$[_{CP_1}$ sein$_i$ **es**   tau esan Mirenek $[_{CP_2}$ t$_i$ jun danik]]
$[_{CP_1}$ Wh$_i$ **NEG** ... $[_{CP_2}$ t$_i$ ...   ]]

(42)   *Clausal pied-piping: intervention by negation*

*$[_{CP_1}$ sein$_i$ t$_{CP2}$ **es**   tau esan Mirenek $[_{CP_2}$ t$_i$ jun danik]]
$[_{CP}$ Wh$_i$ t$_{CP2}$ **NEG** ... $[_{CP_2}$ t$_i$ ...   ]]

While this proposal has many qualities, an analysis in terms of Q-particle makes resorting to reconstruction superfluous. As correctly pointed out by a reviewer, Cable's theory renders pied-piping structures directly interpretable without applying any sort of reconstruction.

6.2.   *Interrogative complements*

Furthermore, another case of restriction alternation arises with complements of interrogative verbs (Ortiz de Urbina 1992, 1993). There are two different kinds of complementizers in Basque: the declarative complementizer *-(e)la* and the interrogative complementizer *-(e)n*. Wh-words in indirect questions selected by the matrix verb may not pied-pipe an embedded interrogative clause (Hualde & Ortiz de Urbina 2003:488). In other words, pied-piping is not allowed in questions with interrogative complements, introduced by verbs like *galdetu* 'ask' and the [+WH] complementizer *-(e)n* (Ortiz de Urbina 1989), as illustrated in (43). Conversely, it is fine with declarative complements introduced by *esan* 'say' and the complementizer *-(e)la* as in (44).

(43)   *[Nor etorriko d-en]   galdetu duzu?
        who   come   AUX-COMP asked   AUX
        'Who will come have you asked?'

(44)   [Nor etorriko d-ela   bihar]   esan diozu Mireni?
        who come   AUX-COMP tomorrow said AUX   Mary.DAT
        'Who will come tomorrow have you told Mary?'   (Ortiz de Urbina 1992:297)

To be more precise, a sentence like (45), which, like (43), is an instance of clausal pied-piping of an interrogative complement (with fronting/topicalization of the subject *Miren*), is grammatical if it takes a yes/no question reading, but not with a matrix wh-reading, indicating a scope effect.

(45)   Mirenek   zer egin zu-en   galdetu duzue?
        Miren.ERG what do AUX-COMP ask   AUX
        *'What did you ask Miren did?'
        'Did you(pl) ask what Miren did?'   (Raynaud 2016:23)

Pied-piping of an wh-interrogative clause selected by a [+WH]-complementizer is permitted when the scope of the interrogative element is restricted to the embedded clause - it can not take matrix scope, as would be expected in (45). That is to say that in the sentence above, the question no longer bears on the wh-word, but it is now interpreted as a yes/no question on the matrix verb *galdetu* 'ask'. The wh-word can only take matrix scope in a pied-piped construction if the verb selects for a [-WH]-complement (like *esan* 'say'), as shown above in (44).

So it looks like in these constructions that involve negation or an interrogative verb, in order to get matrix scope, clausal pied-piping may not occur. Providing we assume [+WH]-verbs to be scope sensitive elements, this would suggest some kind of intervention effect. Although in both cases, the ordering of scope taking elements is the same, i.e. the wh-word c-commands the negation or the [+WH]-verb (wh-word > NEG/[+WH]-verb). What changes is that the negation or the [+WH]-verb do not scope over the embedded sentence anymore in cases of pied-piping.

(46)　　a.　*Wh-extraction*
　　　　　　　$[_{CP1}$ $[_{QP}$ wh-word] ... NEG/[+WH]-verb ... $[_{CP_2}$ ...]]
　　　　　　　QP [wh-word] > NEG/[+WH]-verb > embedded predicate

　　　　b.　*Pied-piping*
　　　　　　　$[_{CP1}$ $[_{QP}$ CP_2] ... NEG/[+WH]-verb ...]
　　　　　　　QP [wh-word > embedded predicate] > NEG/[+WH]-verb

This seems to have the effect of disallowing a matrix scope reading in cases of interrogative complements, while still allowing pied-piping under a yes-no reading question (the interrogative verb no longer has scope over the embedded predicate). In the case of negative sentences, pied-piping seems to be disallowed completely, as negation cannot take scope over the embedded predicate. The relative order of scope-taking elements over the embedded clause seems to make the correct prediction for the unavailability of clausal pied-piping in two contexts, matrix negation and interrogative complements, that have not been previously connected. Moreover, it could account for this class of restrictions in a way that is compatible with a Q-based analysis. Further investigations need to be conducted in order to determine if this proposal could be the right one.

7.　*Conclusion*

The main focus of this paper has been to account for the distribution and the free variation patterns of clausal pied-piping and long-distance wh-movement in Basque wh-questions. In what precedes, I established a more comprehensive picture of Basque wh-questions that substantiates the claims for true optionality. Clausal pied-piping constructions can alternate with their long-distance extraction counterparts without any interpretative effect in many contexts. Based on an analysis of wh-movement that involves a relation between a C head and a QP (Cable 2010b), I have been able to show that the syntactic representations of these constructions are such that they share a same lexical array but differ in the stage of the derivation in which the Q-particle merges. Relying on Biberauer and Richards's (2006) analysis legitimating the possibility of true optionality, I have shown that the proposed Q-based derivations equally satisfy formal requirements, albeit in different ways. Such an analysis can furthermore account independently for several restrictions governing the distribution of clausal pied-piping in Basque, such as the leftness requirement for the pied-piper and constrained pied-piping in cases where the wh-word orginates in an island. Finally, I have attempted to provide an analysis that can explain the scope interactions and restrictions on pied-piping that surface with negation and interrogative complements. Although syntactic optionality remains a debatable issue, this paper would hopefully have made a contribution to the debate by documenting an instance in which syntactic optionality can be accounted for in a Minimalist framework.

Louise Raynaud
Universität Göttingen
louise.raynaud@stud.uni-goettingen.de
Website

References

Adger, D. & J. Smith (2005). Variation and the minimalist program. Cornips, L. & K. P. Corrigan (eds.), *Syntax and Variation: Reconciling the Biological and the Social*, Benjamins, Amsterdam, pp. 149–178.

Adli, A. (2006). French wh-in-situ questions and syntactic optionality: evidence from three data types. *Zeitschrift für Sprachwissenschaft* 2:25, pp. 163–203.

Arregi, K. (2003). Clausal pied-piping. *Natural Language Semantics* 2:11, pp. 115–143.

Beck, S. (1996). Wh-constructions and transparent logical form. PhD dissertation, Universitaet Tuebingen.

Biberauer, M. T. (2003). Verb second (V2) in Afrikaans: a minimalist investigation of word order variation. PhD Dissertation, University of Cambridge.

Biberauer, M. T. & M. Richards (2006). True optionality: when the grammar doesn't mind. Boeckx, C. (ed.), *Minimalist Essays*, John Benjamins, Amsterdam, pp. 35–67.

Cable, S. (2010a). Against the evidence of pied-piping: evidence from Tlingit. *Linguistic Inquiry* 4:41, pp. 563–594.

Cable, S. (2010b). *The grammar of Q/ Q-particles, wh-movement, and pied-piping*. Oxford University Press, Oxford/New York.

Cable, S. (2012). Pied-piping: introducing two recent approaches. *Language and Linguistics Compass* 12:6, pp. 816–832.

Cable, S. (2013). Pied-piping: comparing two recent approaches. *Language and Linguistics Compass* 2:7, pp. 123–140.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge, Mass.

Chomsky, N. (2001). Derivation by phase. Kenstowicz, M. (ed.), *Ken Hale: a life in language*, MIT Press, Cambridge, MA, pp. 1–52.

Cole, P. & G. Hermon (1981). Subjecthood and islandhood: evidence from Quechua. *Linguistic Inquiry* :12, pp. 1–30.

Duguine, M. & A. Irurtzun (2014). From obligatory wh-movement to optional wh-in-situ in Labourdin Basque. *Language* 1:90, pp. 1–30.

Heck, F. (2008). *On pied-piping: wh-movement and beyond*. Mouton de Gruyter, Berlin.

Heim, I. (1992). Kroch, cinque, negative islands. Handout, MIT.

Horvath, J. (2006). Pied-piping. Everaert, M. & H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, Blackwell Publishing, chap. 50, pp. 569–630.

Hualde, J. I. & J. Ortiz de Urbina (2003). *A grammar of Basque*. Mouton de Gruyter, Berlin/New York.

Müller, G. (2001). Optionality in optimality-theoretic syntax. Cheng, L. & R. Sybesma (eds.), *The Second Glot International State-of-the-Article Book*, Mouton de Gruyter, Berlin, pp. 289–321.

Oiry, M. (2011). A case of true optionality: Wh-in-situ patterns likes long movement in French. *Linguistic Inquiry* 1-2:37, pp. 115–142.

Ortiz de Urbina, J. (1989). *Parameters of the grammar of Basque: a GB approach to Basque Syntax*. Foris, Dordrecht.

Ortiz de Urbina, J. (1992). Interrogative discharge and the wh-criterion in basque. *ASJUren Gehirriak: Syntactic theory and Basque syntax* :27, pp. 295–308.

Ortiz de Urbina, J. (1993). Feature percolation and clausal pied-piping. Hualde, J. I. & J. Ortiz de Urbina (eds.), *Generative Studies in Basque Linguistics*, John Benjamins, Amsterdam, pp. 193–208.

Raynaud, L. (2016). Basque wh-questions and syntactic optionality. MA Thesis, University of Cambridge.

Richards, N. (1997). What moves where when in which language. PhD Dissertation, MIT.

Ross, J. R. (1967). Constraints on variables in syntax. PhD Dissertation, MIT.

Simpson, A. & T. Bhattacharya (2003). Obligatory overt wh-movement in a wh-in-situ language. *Linguistic Inquiry* 1:34, pp. 127–142.

The syntax of [ung 'yes'/ani 'no', XP] as right dislocation

Ui-Jong Shin[*] and Sunjoo Choi[**]

The aim of this paper is to investigate how sentences consisting of the polarity answer particle (PAP), *ung* 'yes' or *ani* 'no', in Korean sequentially followed by a fragmentary XP remnant (alias UAX) are considered as one type of a right dislocation (RD) construction. We may misunderstand PAP as just a particle but it is, in fact, a remnant derived through the covert TP ellipsis. We are going to examine four types of UAX-RDs and the parallelism between UAX-RDs and canonical RDs. Both of them show the similar pattern related to the island effect, the full host-clause requirement, voice and verbal morphology matching, and specificational coordination.

1. Introduction

In this paper, we suggest that, the fragmentary sentence which is the polarity answer particle (PAP) *ung* 'yes' or *ani* 'no' plus XP remnant combination is regarded as one type of right dislocation (RD) construction. In previous studies, Ko (2014) and Ahn and Cho (2015, 2016) consider this sentence as a fragment construction. The following example, (1), shows that the question-answer pair includes the positive PAP *ung* 'yes' followed by a fragmentary XP remnant (Ko 2014). The remnant is an answer to an information-seeking question.

(1) Q: Mary-ka   motwu ta   an   manna-ss-ni?
        M.-NOM   all   not   meet-PAST-Q
        (lit.) 'Didn't Mary meet all/any of them?'   (all>>Neg, Neg>>all)

     A: Ung. motwu ta.
        Yes. all (of them)
        (lit.) 'Yes, Mary did not meet all of them.'   (all>>Neg, *Neg>>all)
        (='No, Mary didn't meet any of them.')   Ko (2014: 293, (26))

---

[*]   First Author
[**]   Corresponding Author

In addition, Ahn and Cho (2015) offered four more examples to suggest that the RD construction in Korean is similar to fragment one in related to the syntactic behaviors of island sensitivity in (2) and (3), case-marking in (4) and contrastive focus in (5).

(2)   Absence of correlates & Island sensitivity of fragments
       Q: Cheli-nun   [____ sacwu-n]   mokkeli-lul   peli-ess-ni?
          C.-TOP   bought-RC necklace-ACC throw away-PAST-Q
          'Did Cheli throw away the necklace that the person bought for him?

       A: *Ung,   emma-(ka)
          Yes   Mommy-NOM
           'Yes, Mommy.'   Ahn and Cho (2015: 432, (13))

 (3)   Fragment answer out of complex NPs
       Q: Ne-nun   [____ emma-uy   cha-lul] pilli-ess-ni?
          You-TOP   mother-GEN   car- ACC borrow-PAST-Q
          'Did you borrow mother's car?'

       A: *Ung,   Yenghi-uy.
          Yes,   Y.-GEN
          'Yes, I borrowed Yenghi's mother car.'   Ahn and Cho (2015: 433, (16))

(4)   Obligatory Case marking in fragments - Case of genitive
       Q: Yenghi-ka emma-lul   mannass-tay?
          Y.-NOM   mother-ACC   meet-PAST-Q
          'Whose mother did Yenghi meet?'

       A: Ung, Cheli-uy./*Cheli.
          Yes, C.-GEN/C.
           'Yes, Cheli's.'   Ahn and Cho (2015: 435 (16))

 (5)   Q: Cheli-ka   ecey   Yuni-lul   manna-ss-ni?
          C.-NOM   yesterday Y.-ACC   meet-PAST-Q
          'Did Cheli meet Yuni Yesterday?'

       A: Ani, nayil.
          No   tomorrow
          'No, tomorrow.'   (Ahn 2012: 106; Ko 2014: 301)

This paper is organized as follows: In section 2, we show evidence for analyzing the particle (PAP) *ung* 'yes' or *ani* 'no' plus XP remnant combination as an RD construction. In addition, we also give a detailed account of derivation of UAX-RD constructions. In section 3, we examine the syntactic behavior of the four more UAX-RD types to explain the parallelism between UAX-RDs and canonical RDs.

2. [Ung 'yes'/Ani 'no'], XP remnant Construction

We mentioned that the sentence consisting of a PAP plus XP remnant is treated as one type of RD constructions. The PAP in the UAX-RD construction is not just a particle but a remnant from the covert structure of preceding question phrase before TP ellipsis. It is essential to license the sequentially following XP remnant such as (6).

(6)   Q: John-i   mek-ess-ni?
       J.-NOM   eat-PAST-Q
       'Did John eat?'

    A: ani,   amwu-kes-to.
      no,   anything
      'No, anything.'

    A':*ung, amwu-kes-to.
      yes,   anything
      (lit.)'Yes, anything.'

In (6A) and (6A'), although the negative polarity item (NPI) is identically applied to the answers, the first is tolerated while the second is not. Both of them have a difference in terms of licensing of the NPI remnant by the PAP *ung* 'yes' or *ani* 'no.' According to à la Kramer and Rawlins (2009), Chung (2014), Park (2015), and Holmberg (2015), we presume that the PAP *ung* 'yes' or *ani* 'no' is derived by the TP ellipsis. That is, each example has a different acceptability and the underlying structure prior to the TP ellipsis as follows:

(6')   A: ani   [TP①~~John-i   mek-ess-e~~],   amwu-kes-to   [TP②~~John-i   mek-ci **anh**-ass-e~~].
      no   J.-NOM   eat-PST-DECL, anything   J.-NOM eat-NM**not**-PST-DECL

    A':*ung [TP③~~John-i   mek-ess-e~~],   amwu-kes-to   [TP④~~John-i   mek-ess-e~~].
      yes   J.-NOM   eat- PST-DECL, anything   J.-NOM   eat- PST-DECL

Example (6') shows that the right edged NPI remnant *amu-kes-to* 'anything' may or may not be licensed in the covert preceding host clause. It is in line with the canonical RD construction. There is a derivation of (6') in detail; firstly, the antecedent TP (= question clause) is copied to (6'A①). Then the sentence polarity is changed in accordance with PAP *ung* 'yes' or *ani* 'no'. As mentioned above, we assume the PAP to be a remnant from the covert structure of a preceding question phrase before TP ellipsis. The sentence polarity of (6'A①) turns positiveness into negativeness since, in this case, the negative PAP *ani* 'no' disaffirms the antecedent clause (= question clause). In addition, the XP remnant is also derived from covert TP (=(6'A②)) that is copied from (6'A①) which changes the polarity of the sentence. Since the covert TP (=(6'A②) has the negative polarity, the NPI remnant can be licensed by negation. However, in (6'A'), the polarity of the covert TPs in the host clause (=6'A'③) and in the appendix clause (=6'A'④) is not changed because of the positive PAP *ung* 'yes' that does not disaffirm the sentence.

It is to be noted that the negative PAP *ani* 'no' does not always license the NPI remnant. In contrast, the positive PAP *ung* 'yes' may be a licensor for the NPI remnant when the question sentence has a negative polarity one, such as (7).

(7)   Q: John-i   mek-ci   anh-ass-ni?
          J.-NOM   eat-NM   not-PAST-Q
          'Did John eat?'

      A: ung,   amwu-kes-to.
          Yes,   anything
          (lit.) 'Yes, anything.'

      A':*ani,   amwu-kes-to.
          No,   anything
          'No, anything.'

The question sentence (7Q) has a negative expression *an* 'no'; in this case, the answer (7A) is acceptable while (7A') is not. The contrast between (7A) and (7A') reinforces our suggestion that not only overt PAP but also the immediately following covert 'reconstructive' TP can license the NPI remnant. Likewise, the XP remnant following the comma is derived in a similar way as the canonical RDed element.

Another piece of evidence for the analysis of UAX as a remnant to the information-seeking question is related to the type of question phrase (QP) and answer pair. In example (8), we can examine the question and variety of its answers pair as the example that supports our suggestion. The prosodic contour is important in (8); the question sentence is to be a 'yes-no question' when it carries rising intonation. On the contrary, when the accent descends at the sentence final position, the sentence is to be a 'wh-question'.

(8)   Q: John-i   mek-ess-ni ↗ ?[1]
          J.-NOM   eat-PAST-Q
          'Did John eat?'

      A: *ung, mwues-ul?
          yes,   what-ACC

      B: *(John-i)   mek-ess-e ↘,   mwues-ul?
          J.-NOM   eat- PST-DECL, apple-ACC

---

[1] In Korean, the null argument may be either definite or indefinite. When it is indefinite, its identity may be a focus in question, yielding a fragmentary question, as follows:

(i) A: John-i   mek-ess-e ↘.
       J-NOM   eat-PST-DECL
    B: mwues-ul?
       what-ACC

(B)'s question is regarded as a sluiced question.

C: mek-ess-e↗, mwuesul?   ('surprise' echo yes-no, followed by 'what')
   eat- PST-DECL, apple-ACC

D: ung, banana-lul.
   yes, banana-ACC

E: (John-i) mek-ess-e↘, banana-lul.
   J.-NOM   eat- PST-DECL, banana-ACC

As the answer to the question, (8A) and (8B) with a question phrase (QP) remnant are not tolerated while (8D) and (8E) with a non-question phrase remnant are grammatical. In this regard, the PAP *ung* 'yes' or *ani* 'no' which marks the sentence polarity is to be a remnant prior to TP ellipsis or is a TP anaphora. The full clause answers such as (8B) and (8E) is also used instead of PAP remnant, thus, we confirm that both the PAP remnant and the full clause act along the same lines with regard to the sentence involving QP or non-QP remnant.

In addition, we need to examine the sentence with a polarity question including the indefinite phrase *mwues* which has been ambiguous interpreted as 'what or something.'

(9) Q: John-i   mwues-ul   mek-ess-ni ↘/↗ ?
      J.-NOM   what/something-ACC   eat- PAST-Q
      'What did John eat?'

     A: *ung, mwues-ul?   after ↗
       yes,   what-ACC

     B: *(John-i)   mek-ess-e ↘, mwues-ul?   after ↗
       J.-NOM   eat-PST-DECL, what-ACC

     C: banana-lul.   after ↘
       banana-ACC

     D: ung, banana-lul.   after ↗
       yes,   banana-ACC

     E: (John-i) mek-ess-e ↘, banana-lul. [2]   after ↗, OK / ↘, *
       J.-NOM   eat- PST-DECL, banana-ACC

When the sentence carries an ascendable intonation at its final, the indefinite phrase *mwues* 'what or something' is construed as 'something.' (9A) and (9B) are ruled out whereas (9D) and

---

[2] The ambiguously interpreted indefinite in subject position as follows is also answered with the RD sentence when the former carries rising intonation.

(i) A: nwu-ka   banana-lul   mek-ess-ni ↗ ?
      who-Nom   banana-Acc   eat-Pst-Q
      'Who ate an banana?'
   B: banana-lul   mek-ess-e,   John-i.
      banana-Acc   eat-Pst-Decl, J-Nom
      'Ate an banana, John.'

(9E) are ruled in. It means, as we mentioned, that we confirm again the characteristic of parallel usage between the PAP remnant and the full clause.

3. More discussions

Before investigating their syntactic behavior in more detail, let us classify that there exists a four type of UAX-RD in Korean. This classification follows Ko's (2014, 2015, 2016) nomenclature of general RDs.

The first type in (10) is categorized by the presence of a gap in the preceding PAP host clause that answers the gap-containing question clause:

(i) Gapped UAX-RD

(10)   Q: kyoswu-nim-I   [e]   chingchanha-sy-ess-ni?
            Professor-HOR-NOM   praise-HOR-PST-Q
            'Did the professor praise?'

       A: Ung, han   nam-haksayng-ul/   John-lul.
            Yes   one   male student-ACC/   John- ACC
            'Yes, one male student/John.'

As illustrated in (10), it can be answered by either definite or more specificational element since the gap in which the host clause is indefinite.

Let us turn to repetitive UAX-RD as shown in (11). As in (11), the answer clause contains a repetitive element *han nam-haksang-ul*, which repeats its correlate in the preceding PAP host clause. The semantic function of the constituent is to put a strong emphasis on the correlate by repeating it.

(ii)   Repetitive UAX-RD

(11)   Q: Kyoswunim-I   han nam-haksayng-ul chingchanha-sy-ess-ni?
            professor-HOR-NOM one male student-ACC praise-HOR-PST-Q
            'Did the professor praise one male student?'

       A: Ung, han   nam-haksayng-ul.
            Yes   one   male student-ACC
            'Yes, one male student.'

The third type is illustrated in (12), where the remnant shares more specific information for a less informative constituent that is suggested in the preceding polarity question sentence. We expect that the remnant in specificational UAX-RD would necessarily be more specific or informative than the correlate in the question clause.

(iii)   Specificational UAX-RD

(12)   Q: Kyoswy-nim-i   han haksayng-ul chingchanha-sy-ess-ni?
            Professor-HOR-NOM one student-ACC   praise-HOR-PST-Q

'Did the teacher praise one student?'
A: Ung, <u>John-lul</u>.
   Yes   John-ACC
   'Yes, John.'

More specifically, the remnant *John* gives a special value for the correlate *han-haksang-ul* in the preceding clause.

The last type is categorized as contrastive UAX-RD. Starting with the negative PAP, it plays a role of negating one property introduced in the preceding polarity question sentence. Furthermore, it makes room for adding a contrasting alternative to it in favor of the following remnant, e.g., *bill-lul* instead of *John-lul* as in (13).

(iv)   Contrastive UAX-RD

(13) Q: kyoswu-nim-i   <u>John-lul</u>   chingchanha-sy-ess-ni? [3]
        Professor-HOR-NOM John-ACC   praise-HOR-PST-Q
        'Did the professor praise John?'

     A: Ani, <u>Bill-lul</u>.
        No,   Bill-ACC
        'No, Bill.'

We have discussed the variability of UAX-RDs in Korean so far. In the following section, we will investigate the syntactic behavior of UAX-RDs.

3.1. Island (in)sensitivity in UAX-RDs

According to Ahn & Cho (2015), the gapped UAX-RD as in (14) shows the island effects in general. In particular, they claim that the fragment is indeed sensitive to island effects if there is no overt correlate in the context, as illustrated in (14A).

▷ gapped UAX-RD
  (14) Q: Cheli-nun [_ sacwu-n] mokkeli-lul peli-ess-ni?
        Cheli-TOP   bought-RC necklace-ACC throw.away-PST-Q
        'Did Cheli throw away the necklace that the person bought for him?

     A: *Ung, **emma-(ka)**
        Yes   Mom-NOM
        'Yes, (his) Mom (bought the necklace for him.)'

                                  Ahn & Cho (2015: 432)

---

[3] The *canonical RD* counterpart to the contrastive UAX-RD in (13) is (i) below:
(i) sensayng-nim-I   yengho-lul   mann-ass-e,   ani cinho-lul.
   teacher-HOR-NOM   Yengho-ACC meet-PST-DECL, not Cinho-ACC
   'The teacher saw Yengho, well, not him but Cinho.'
The remnant is preceded by the negator *ani* 'not.

In contrast, consider the following examples of gapped UAX-RD, constructed on the model of the sentences that Kuno (1973) used to examine island sensitivity in relativization of Japanese [4]:

(15) a. [[____ᵢ ip-un   os-i]   ccice-cy-ess-e],   ku   haksayng-iᵢ.
         wear-PST-REL   clothes-NOM   torn-be-PST -DECL,   that   student-Nom
         'The clothes ((s)he) wore was torn, that student.'

     b. Q: [[____ᵢ ip-un   os-i]   ccice-cy-ess-ni]?
           wear-PST-REL   clothes-NOM   torn-be-PST-Q
           'Was the clothes ((s)he) wore torn?'

        A: Ung, ku   haksayng-iᵢ.
           Yes, that   student-NOM
           'Yes, that student.'

(16) a. [[__ᵢ   casal-ha-yss-umeyto   pwulkwuhako]   amwuto sulpheha-ci
                 suicide-do-PST-   in spite of   anyone grieve-NM
         anh-ass-e]], ku   paywu-kaᵢ.
         not-PST-DECL that   actor-NOM
         'No one grieved over his suicide though he took a suicide, that actor.'

     b. Q: [[__ᵢ casal-ha-yss-umeyto pwulkwuhako]   amwuto sulpheha-ci   anhass-ni]?
                 suicide-do-PST-in spite of   anyone grieve-NM   not-PST-Q
        A: Ung, ku   paywu-kaᵢ.
           Yes, that   actor-NOM
           'Yes, that actor.'

Even though the remnant here is linked to a gap within a relative clause island, either canonical gapped RDs in (15a) and (16a) or gapped UAX-RDs in (15b) and (16b) are fine. The question that is naturally raised is why some (UAX-)RDs are acceptable, while some others are unacceptable in terms of island effects. We assume that the discourse status of a gap in gapped (UAX-)RDs plays a crucial role when it comes to island effects (cf. Kim and Kuno 2013). For

---

[4] See Han and Kim (2004) for the *double relativization* analysis of island-circumventing extraction. But see also Yoon (2011) for its critique. For example, (i) modeled after one of the examples in Yoon (2011) where double relativization (thereby, subjectivization) out of a complex NP island is not allowed is unacceptable:
(i) *ku   capci-ka kyengchal-I kwutokhaten tokca-lul motwu   cheyphohayssta.
    That magazine the police-NOM subscribed to people-ACC all arrested
    'That magazine. the police arrested the people that subscribed to.'
However, when the example in (i) is comprised of the question and the gapped/repetitive UAX-RD pair as in (ii), we cannot see the island effects (However, the implicit argument (Ø) in this example has to be construed as specific.)
(ii) Q: kyengchal-I   [Ø/capci-lul   kwutokhaten   tokca-lul]   motwu   cheyphohayss-ni?
        The   police-NOM [magazine-ACC   subscribed   to people-ACC] all   arrested-Q
        'Did the police arrest the people that subscribed to (the magazine)?'
     A: ung,   capci-lul/sasangkyey-lul.
        yes,   magazine-ACC/-ACC
        'Yes, a magazine/Sasangkyey'

The grammaticality of (ii) can be a crucial evidence that double relativization is not always responsible for island sensitivity.

this reason, we suggest that there are two types of gap in the host clause of (UAX-)RDs. In other words, one is d-linked *pro* which shows island effects, and the other is non-d-linked *pro* which does not represent island effects.

Let us consider gapless UAX-RDs. Ahn and Cho (2016) show that repetitive UAX-RDs, as in (17), does not show island effects. Moreover, Ko (2016) note that specificational UAX-RDs as in (18) are indeed island sensitive:

▷ gapless   UAX-RD

(17) Q: Cheli-nun　　　　　[emma-ka　sacwu-n]　mokkeli-lul　peli-ess-ni?
　　　　　C.-TOP　　　　　　mom-NOM　bought-RC　necklace-ACC　throw.away-PST-Q
　　　　　'Did Cheli throw away the necklace that the person bought for him?

　　　A: Ung, **Emma-(ka)**
　　　　　Yes　Mommy-NOM
　　　　　'Yes, Mommy'　　　　　　　　　　　　　　　　　　Ahn and Cho (2015:432)

(18) Q: Cheli-　ka kacok-I　　sacwu-n cha-lul　ilhepeli-ess-ni?
　　　　　C.-NOM　family-NOM　buy-RC　car-ACC　lose-PST-Q
　　　　　'Did Cheli lose the car that his mother bought for him?'

　　　A: *Ung,　emma-ka.
　　　　　Yes,　mom-NOM 'Yes, (his) mom (bought the car for him).'
　　　　　　　　　　　　　　　　　　　　　　　　　　　　　Ko (2016: 10)

We disagree with Ko (2016)[5]' analysis but adopt Ahn and Cho (2016), saying that neither repetitive nor specificational/contrastive (UAX-)RDs are island sensitive[6], as shown in (19)-(22)[7]:

(19) a. kyoswu-nun <u>haksayng-I</u> molko o-n cha-lul po-ass-e,
   professor-TOP student-NOM drive-and come-REL car- ACC see-PST-DECL
   {haksayng-i, john-i} ~~molko~~ ~~o-n~~ ~~cha-lul~~
   student-NOM J-NOM drive-and come-REL car- ACC
   'The professor saw the car that a student had driven there, {a student, John}.'

  b. kyoswunim-un [John-i <u>chayk-ul</u> pilly-e cwu-n haksayng-ul] chacko iss-e,
   professor-TOP J-NOM book- ACC borrow give-REL student-ACC looking beDEC
   [~~John-i~~ chayk-ul/ coy-wa pel-ul ~~pilly-e cwu-n~~ ~~haksayng-ul~~].
   J-NOM book-ACC /crime-and punishment-ACC borrow give-RELstudent-ACC
   'The professor is looking for a student who lent a book to, a book/Crime and Punishment.'

(20) Q: John-i <u>yenghwa-lul</u> po-n chinkwu-ul chacko iss-ni?
   John-NOM movie-ACC watch-REL friend-ACC look-Progress-Q
   'Is John looking for a friend that watched a movie?'

  A: Ung, <u>yenghwa-lul</u>/<u>silmito-lul</u> (~~po-n salam-ul~~).
   Yes movie-ACC /Silmito-ACC watch-REL person-ACC

---

[5] We agree with Ko (2016) on the unacceptability of (18), noting that the noun *kacok* 'family' does not comport well with the specificational remnant like *emma*. See (19)-(22) for other nouns that are agreeable with their corresponding 'specificational' remnants.

[6] Ahn and Cho (2016) note that the repetitive type of UAX-RD display island effects when the remnant is accompanied by a sentential adverb as in (i):
(i) A: Cheli-nun [emma-ka sacwu-n] mokkeli-lul peli-ess-ni?
  C.-TOP mom-NOM bought-REL necklace-ACC throw.away-PST-Q
  'Did Cheli throw away the necklace that the person bought for him?'
 B: *Ung, pwunmyenghi emma-ka.
  Yes clearly mom- NOM
  'Yes, clearly mom.'
Unlike the repetitive type of UAX-RD, however, the contrastive type of UAX-RD does not show such effects, as in (ii):
(ii) Q: sensayngnim-un cheli-ka <chayk-ul> pillyecwu-n haksayng-ul chacko iss-usi-ni?
  teacher-TOP Cheli-NOM book-ACC lent-REL student-ACC look be-HOR-Q
  'Is the teacher looking for the student that Cheli lent a book to?'
 A: ?ani, amato khemphyuthe-lul.
  No maybe computer- ACC
  'No, maybe a computer.'
As discussed in the footnote 8, what is at stake in distinguishing repetitive and contrastive UAX-RDs in accompanying a sentential adverb is the focus relation between the remnant and its correlate. We suspect that the more prominent relation between them in contrastive UAX-RDs improve on the acceptability of the sentential adverb-accompanied remnant that is associated with its correlate inside an island structure.

[7] We acknowledge that there is a speaker variation on the acceptability of (19)-(22). As far as we can see, the most important factor that improves on their acceptability seems to be the explicit/implicit focus on the correlate in the host clause that is interpretively linked to the corresponding remnant.

'Yes, a movie/Silmito.'

(21) Q: Kyoswunim-un [John-i <u>chayk-ul</u> pilly-e cwu-n haksayng-ul] chacko iss-ni?
         Professor-TOP   John-NOM book-ACC borrow give-RELstudent- ACC   looking be-Q

      A: Ani, [~~John-i~~   khemphyuthe-ul ~~pilly-e cwu-n   haksayng-ul~~].
        No, John-NOM computer-ACC   borrow give-REL   student- ACC
        'No, a computer.'

(22) Q: Kyoswunim-un [John-i <u>chinkwu-lopwuthe</u> pat-un   chayk-ul] kemsahayss-ni?
         Professor-TOP   Johni-NOM friend-from   receive-REL book- ACC inspected-Q
        'Did the professor inspect a book that John received from her friend.'

      A: Ani, [~~John-i~~   eni-lopwuthe ~~pat-un   chayk-ul~~]
        No, Johni-NOM sister-from   receive-REL book- ACC
        'No, from her sister.'

We attribute the island insensitivity of gapless (UAX-)RDs in Korean due to Optional Max-Elide to this language that strips off the right-edge element(s) of island-forming structure, with the help of clausal ellipsis (Park, Shin and Kang 2015; An (2016)). In so doing, it is ultimately yielding the surface remnant without causing a violation of any island constraint. This can be understood as 'peripheral ellipsis' (ellipsis that occurs at the periphery of a clause) named Max-Elide (following the similar concept of the operations *Max-Elide*) in Merchant (2008)), which applies in the following mode:

(23) Optional Max-Elide
      Max-Elide applies not only to the TP where remnants/ survivors move out of, but also optionally to the right periphery of the final remnant/survivor immediately preceding the elided TP.

In other words, the right-edge element(s) of island-forming structure as part of the alternatively larger remnant is included through Optional Max-Elide in the portion that undergoes clausal ellipsis. As a result, in Korean RDs including UAX-RDs there is no difference between NP and non-NP remnants (cf. Hoji 1990), nor any difference between informational and contrastive focus remnants (cf. Griffith and Liptak 2014).

    Let us consider the structure that is subject to the left branch condition. As noted by Ahn and Cho (2016), the Genetive-marked remnant is possible when it has its correlate in the host clause of specificational UAX-RD as shown in (23).

(23) LBC in UAX-RD
      Q: Yenghi-ka wulipan namca-uy emma-lul   manna-ss-ni?
         Y.-NOM our   class boy-GEN   mother-ACC   meet-PST-Q
        'Did Yenghi meet the mother of a boy in our class?'

      A: Ung, <u>Cheli-uy</u>./(*)Cheli.
        Yes, C.-GEN/C.
        'Yes, (Yenghi met) Cheli's (mother).'

                                        (Ahn and Cho 2016: note 9)

However, as illustrated in (24), Ko (2013) said that it does not relate with its correlate preceding another Genitive-marked element.

(24) Q: Ney-ka   chinkwu-(uy) emma-uy   cha-lul   pilli-ess-ni?
          You-NOM   friend-(GEN) mother-GEN   car- ACC borrow-PST-Q
          'Did you borrow a friend's mother's car?'

      A:*Ung,   Yenghi-(uy).   <===*   because of intervention effects
         Yes,   Y-GEN
         'Yes, (I borrowed) Yenghi's (mother's car).'   (Ko 2016: 11)

Due to the intervention effects induced by the intervening identically Genitive-marked element that occurs between Genitive-marked remnant and its corresponding Genitive-marked correlate, the unacceptability of (24A) can be observed (cf. Ko 2014, 2015; D. Chung 2015)[8] . By contrast, (23A) where there is no such kind of intervening element is fine. Incidentally, Ahn and Cho (2016) claim that the remnant in (23A) is required to carry the same Case as its correlate in the host clause. However, we are on the different track that the Genitive Case can be optional.

Crucially, however, we agree with Ahn and Cho (2016)'s analysis, suggesting that either specificational or contrastive UAX-RDs permit the remnant that is obviously derived from the left branch structure, as in (25)-(26):

(25) Q: John-i   wulipan   yehaksayng(-uy)   nothupwuk-ul swuliha-yess-ni?
          John-NOM our   class female student-GEN   notebook-ACC repair-PST-Q
          'Did John repair our female classmate's notebook?'

      A: Ung, Mary-uy/?Mary (nothupwuk-ul)
         Yes, Mary-GEN   notebook-ACC
         'Yes, Mary('s)'.

(26) Q: John-i   wulipan namhaksayng(-uy)   nothupwuk-ul hwumchy-ess-ni?
          John-NOM our   class male student-GEN   notebook-ACC steal-PST-Q
          'Did John steal our male classmate's notebook?'

      A: Ani, wulipan yehaksayng(-uy)   (nothupwuk-ul)
         No   our   class female student-GEN   notebook-ACC
         'No, our female classmate('s)'.

---

[8] These invention effects seem to be suspended when the intervening element is of a different (semantic) import from both the remnant and its correlate, as follows:
(i) Q: Chelswu-ka nay   chinkwu-uy yeph   cip-uy taymwun-ul pal-lo   chass-ni?
        Chelswu-NOM   my friend-GEN neighbor-GEN gate-door-ACC   foot-with kick-Q
        'Chelswu kicked my friend neighbor's neighbor's gate-door.'
    A: Ung, yenghuy-uy.
       Yes   Yenghuy-GEN
       'Yes, Yenghuy's.'

What is crucial here is the assumption that Optional Max-Elide applies phonologically suppressing the right-edge element(s) of the remnant, there is no derivational point where the LBC constraints apply, safely explaining these examples of (25)-(26).

3.2. The 'full' host clause requirement in UAX-RDs

Ahn and Cho (2016) further examine the contrast between (27) and (28). Observing the grammaticality of (27), they also mention that the remnant fragment of (27A) in their analysis cannot meet the identity condition on ellipsis with the preceding question clause in (27Q).

(27) Q: Cheli-ka Yenghi-lul cala-ss-ni?
        C.-NOM   Y.-ACC   cut-PST-Q
        'Did Cheli cut Yenghi?'   (Ahn and Cho 2016, (44))

      A: *Ung,   meli-lul
        Yes,   hair-ACC
        'Yes, (he cut Yenghi's) hair.'

However, they assume that the remnant fragment of (28A) can satisfy the identity condition on ellipsis with the preceding question clause in (28Q).

(28) Q: Cheli-ka Yenghi-lul ttayli-ess-ni?
        C.-NOM   Y-ACC   hit-PST-Q
        'Did Cheli hit Yenghi?'

      A: Ung, meli-lul.
        Yes, head-ACC
        'Yes, (he hit Yenghi's) head.'   (Ahn and Cho 2016, 45)

Contra Ahn and Cho (2016)[9], we claim that the contrast between (27) and (28) blamed on what Kuno (1978) calls the 'full' host clause requirement. It allows that the host clause before the RD-ed remnant must be syntactically/semantically full or complete. In other words, the unacceptability of the question-answer pair in (27) is owing to the violation of the selection restriction on the part of (27Q). A person *Yenghi-lul* cannot be a target that can be cut by another person in general.
    The same point can be observed. The following examples make the point:

(29) Q: John-i   Bill-lul   <u>mit</u>-ni?

---

[9] When the selectional restriction using the same verb is met, the sentence improves as follows:
(i) Q: Chelswu-ka   sayngsen-ul   call-ass-ni?
        Chelswu-NOM   fish-ACC   cut-PST-Q
        'Did Chelswu cut fish?'
    A: Ung, meli-lul.
        Yes   head-ACC
        'Yes, a head.'

John-NOM   Bill-ACC believe-Q
'Does John believe Bill?'

    A: ?ung, ttokttokha-ta-ko.
       yes, smart-DECL-COMP
      'Yes, as smart.'

(30) Q: John-i   Bill-lul   <u>sayngkakha</u>-ni?
      John-NOM   Bill-ACC   think-Q
      'Does John think Bill?'

    A: ???ung, ttokttokha-ta-ko.
        Yes,   smart-DECL-COMP
      'Yes, as smart.'

Note here that the examples (31) and (32) are a minimal pair, like those in (27) and (28); the only difference is the verbs. According to Park and Shin (2015), the verb *mit-* 'believe' can undergo syntactic reanalysis when encountering the RD-ed element, changing from an NP/DP-subcategorizing verb to a CP subcategorizing verb because this change is made on the comparable semantic dimension. However, the verb *sayngkakha-* 'think' cannot do so due to the corresponding reanalysis involving a radical change from one semantic dimension to another semantic dimension of meaning[10].

    In contrast, the negative PAP cannot be followed by the predicate remnant for the verb *mit-* 'believe', needless to say the verb *sayngkakha-* 'think', as illustrated in (31) and (32):

(31) Q: John-i   Bill-lul mit-ni?
      John-NOM   Bill-ACC believe-Q
      'Does John believe Bill?'

    A: *ani, papo-la-ko.
      No,   fool-COP-COMP
      'No, as a fool.'

(32) Q: John-i   Bill-lul sayngkakha-ni?
      John-NOM   Bill-ACC think-Q
      'Does John think Bill?'

---

[10] The verb *sayngkakha-* 'think' and the verb *mit-* 'believe' have the following meanings defined in 「The Standard Korean Great Dictionary published by the National Institute for the Korean Language」. The (a)-meanings hold when they take an NP/DP complement, whereas The (b)-meanings hold when they take a clausal complement.
(i) *mitta* 'believe'
   a. have a confidence in someone and take him/her as reliable
   b. take someone or something as ∽ without any doubt
(ii) saygkakhata 'think'
   a-(i). remember about someone
   a-(ii). show one's good faith to/give one's whole mind to someone
   b. have an opinion or feeling about something

A: *ani, ttokttokha-ta-ko.
   No,   smart-DECL-COMP
   'No, as smart.'

The grammaticality of these examples indicate that the negative PAP presupposes an overt correlate in the preceding question clause that corresponds to the RD-ed remnant. The absence of an overt correlate in the preceding question clause is responsible for the grammaticality of (31A) and (32A). An account based on the presence or absence of its overt correlate makes a correct prediction.

   The examples in (33) and (34) are also compatible with the analysis as well. This can be analyzed in the same way.

(33) Q: Ne   John-lul   po-ass-ni?
       You John-ACC   see-PST-Q
       'Did you see John?'

    A: (?)ung, wuntongha-nun   kes-ul.
       Yes,   exercise-REL   thing-ACC
       'Yes, doing an exercise.'

(34) Q: *Ne John-lul   kancwuha-ni/yeki-ni?
       You John-ACC   regard-Q/take-Q
       'Do you regard/take John?'

    A: ung, chencay-la-ko.
       Yes, genius-COP-COMP
       'Yes, as a genius.'

The verb *po-* 'see' in (33) may change from an NP-taking to a nominal clause-taking verb because its change is made on the same semantic dimension, making the acceptable UAX-RD. The example (34), however, is excluded because neither the verb *kancwu-* 'regard' nor *yeki-* 'take' can be used as an NP-taking verb.

   Regarding to the 'full' host clause requirement is head restriction, which shows that adnominal modifiers can undergo RD, but the head (projections) of NP leaving their modifiers behind in the host clause cannot (cf. Simon 1989; Im and Chang 1995; C-H Lee 2009; Yoon 2013, etc.). The examples in (35) show head restriction. The materials are taken from Yoon (2013):

(35) a. na-nun [____   Mary]-ul   mannass-e, wuli-thim-ey saylo   haplyuha-n
      I-TOP   Mary-ACC met-DECL   our-team-LOC newly   joined-REL
      'I met Mary, who recently joined our team.'
   a'. na-nun [wuli-thim-ey saylo   haplyuha-n____] mannass-e,   [Mary]-ul.
      I-TOP   our-team-LOC newly   joined-REL   met-DECL,   Mary-ACC
   b. na-nun[____   cha]-lul   pilli-ess-e,   cinan pen-kwa   ttokkathu-n.
      I-TOP   car-ACC   borrow-PST-DECL   last time-as   same-REL
      'I borrowed the same kind of car as last time.'

b'. *na-nun [cinan pen-kwattokkathu-n___] pilli-ess-e,   cha-lul
   I-TOP last time-as same-REL   borrow-PST-DECL car-ACC

Leaving only their modifiers behind in the host clause, the host clause cannot be 'full' or complete when the head (projections) of NP goes missing. By contrast, the syntactically option modifiers may be derived later as a RD-ed element without violating any grammatical constraint. To be also included in the realm covered by the 'full' host clause requirement is predicate restriction observed by Chung (2009, 2011), which rules that while arguments/adjuncts can undergo RD, predicates cannot. The following examples exemplify predicate restriction:

(36) a. Bill-un [nay-ka John-lul cikcep mannassta-ko] sayngkakhay.
    Bill-TOP I-NOM John-ACC in.person met-COMP   thinks.
    'Bill thinks that I met John in person.'
   b. Bill-un [nay-ka __ cikcep   mannassta-ko] sayngkakhay, <u>John-lul</u>.
    Bill-TOP I-NOM in.person met-COMP thinks   John-ACC
   c. *Bill-un [nay-ka John-lul cikcep __] sayngkakhay, <u>mannassta-ko</u>.
    Bill-TOP I-NOM John-ACC in.person thinks   met-COMP
   d. *Bill-un [nay-ka John-lul__] sayngkakhay, [<u>cikcep   mannassta-ko</u>].
    Bill-TOP I-NOM John-ACC thinks   in.person met-COMP
   e. * Bill-un [nay-ka __] sayngkakhay, [<u>John-lul cikcep   mannassta-ko</u>].
    Bill-TOP I-NOM   thinks   John-ACC in.person met-COMP
   f. Bill-un _ sayngkakhay, [<u>nay-ka John-lul cikcep   mannassta-ko</u>].
    Bill-TOP thinks   I-TOP John-ACC in.person met-COMP

As mentioned above, the host clauses in (36c-e) independently cannot count as full or complete. The data also show that the (36c-3) can be analyzed in the same way as in the contrast between (27) and (28). This prediction is safely borne out as well.

3.2. Case or voice match in UAX-RDs

Note that canonical RDs in (37) require Case/voice match. The data presented in (37) show that this is indeed the case.

(37) a. ecey   John-i namwu-lul call-ass-e, namwu-lul/*namwu-ka.
    Yesterday John-NOM tree-ACC cut-PST-DECL tree-ACC/tree-NOM
    'Yesterday John cut, a tree.'

   b. ecey   namwu-ka John-ey uyhay cal-ly-ess-e, namwu-ka/*namwu-lul.
    Yesterday tree-NOM John-by cut-PASS-PST-DECL tree-NOM /tree-ACC
    'Yesterday a tree was cut by John, a tree.'

As in (37), the RD-ed element and its correlate must match in terms of Case. In addition, if the RD-ed remnant is derived from a clausal structure, the RD clause and its corresponding host clause must match in voice.

 Given the discussion so far, we can straightforwardly extend this analysis for UAX-RDs. As expected, UAX-RDs in Korean display the same case/voice match effects as shown in (38) and (39):

(38) Q: ecey   John-i   <u>namwu-lul</u> call-ass-ni?
      Yesterday John-NOM tree-ACC   cut-PST-Q
      'Did John cut a tree yesterday?'

   A: ung, <u>namwu-lul</u>/*<u>namwu-ka</u>.
      Yes tree-ACC/   tree-NOM
      'Yes, a tree.'

(39) Q: ecey   namwu-ka   John-ey uyhay   cal-ly-ess-e,   namwu-ka/*namwu-lul.
      Yesterday tree-NOM   John-by   cut-PASS-PST-DECL, tree-NOM/tree-ACC
      'Yesterday it was cut by John, a tree.'

   A: Ani, <u>tol-i</u>/ *<u>tol-ul</u>.
      No   a stone-NOM/a stone-ACC
      'No, a stone.'

However, (40) and (41) where the complex verb such as *sa-ko sip-* 'buy-C want' allows Case alternations of accusative or nominative for the preceding object NP, are acceptable even when the RD remnant mismatches in Case form with its correlate in the host clause, as pointed out by Ahn and Cho (2016)[11]:

(40) Q: Ne-nun   cha-lul sa-ko   sip-ni?
      You-TOP   car-ACC buy-COMP   want-Q
      'Do you want to buy a car?'

   A: Ung, BMW-ka.
      Yes   BMW-NOM
      'Yes, (I want to buy) a BMW.'

(41) Q: Ne-nun   cha-ka   sa-ko   sip-ni?
      You-TOP   car-NOM   buy-COMP   want- Q
      'Do you want to buy a car?'

   A: Ung, BMW-lul.
      Yes, BMW-ACC.
      'Yes, (I want to buy) a BMW.'

The acceptability of (40) and (41) indicate that when it comes to the identity condition on ellipsis, we rather consider such elements as voice morphology (cf. Merchant 2013), not Case identity.

3.4. Specificational coordination

---

[11] The examples in (40) and (41) as well as (42), (43), and (45) are taken from the earlier version of Ahn and Cho (2016).

Ko (2016) and Ahn and Cho (2016) note that the remnant in UAX-RDs should be more informative than its correlate in the preceding question clause. In the previous section, we have shown that the PAP is used to affirm or deny the preceding question clause, so that the PAP itself does not provide specific informative content for the correlate in the question clause. Instead, the element after the following comma provides a specific value.

(42)   Q: Yenghi-ka <u>ku haksayng-ul</u>   manna-ss-ni?
          Y.-NOM   the student-ACC   meet-PST-Q
          'Did Yenghi meet the student?'

       A: Ung, <u>Cheli-lul</u>.
          Yes, C.-ACC.
          'Yes, (Yenghi met) Cheli.'

(43)   Q: Yenghi-ka <u>haksayng-ul</u>   manna-ss-ni?
          Y.-NOM   student-ACC   meet-PST-Q
          'Did Yenghi meet a student?'

       A: Ung, <u>Cheli-lul</u>.
          Yes, C-ACC.
          'Yes, (Yenghi met) Cheli.'

(44)   Q: Cheli-ka <u>haksayng-ul</u>   ttayli-ess-ni?
          C.-NOM   student-ACC   hit-PST-Q
          'Did Cheli hit a student?'

       A: Ung, <u>SNU-haksayng-ul</u>.   A': #Ung,   <u>salam-ul</u>.
          Yes, SNU-student-ACC   A': Yes,   person-ACC
          'Yes, (he hit) an SNU student.' 'Yes, (he hit) a person.'   (Ko 2016: note 5)

(45)   Q: John-i   achim-ul   <u>kuke-lul</u> mek-ess-ni?
          J.-NOM   breakfast-ACC it-ACC   eat-PST-Q
          '(lit.) Did John eat breakfast, it?'

       A: Ung, <u>banana-lul</u>.
          Yes, banana-ACC
          'Yes, (John) ate a banana.'

The use of the PAP is noteworthy. The positive PAP is used to affirm the previous question, and is followed by the *NP modifier* remnant as in (46A). However, the negative PAP negates the proposition depicted by preceding question (thereby denying the existence of the NP to be modified by the *NP modifier* remnant). In doing so, the *NP modifier* remnant added posterior to the denied proposition seems unnatural, as shown in (46A'):

(46)   Q: ku-ka   han   sonyen-ul manna-ss-ni?
          he-NOM one   boy-ACC   meet-PST-Q

'Did he see a boy?'

    A: ung, [acwu   ttokttokha-ko calsayngki-n]
        yes,   very   smart-and   handsome-REL
        'Yes, very smart and handsome.'

    A':#ani, [acwu   ttokttokha-ko calsayngkin].
        no,   very   smart-and   handsome-REL
        'No, very smart and handsome.'

Note that the negative PAP plays another crucial role. As pointed out above, it negates not only the preceding full clause but also one of the constituents in it, as exemplified in (47) and (48):

(47) Q: John-i   [<u>khi-ka   acwu khu-n</u>] sonyen-ul manna-ss-ni?
        John-NOM height-NOM   very tall-REL boy-ACC   meet-PST-Q
        'Did John see a very tall boy?'

    A: ung, [khi-ka   acwu khu-n]
        Yes height-NOM   very tall-REL
        'Yes, very tall in height.'

    A':ani, [khi-ka   acwu   caku-n].
        No   height-NOM very   small-REL
        'No, very small in height.'

(48) Q: ku-ka   John-<u>uy</u>   tongsayng-ul manna-ss-e.
        he-NOM John-GEN brother-ACC   meet-PST-Q
        'Did he see John's brother?'

    A: ung, John-uy.
        yes, John-GEN
        'Yes, John's.'

    A': ani, Bill-uy.
        No   Bill-GEN
        'No, Bill's.'

As is obvious by now, it naturally follows that the negative PAP (as the constituent negation) demands the presence of a constituent in the preceding question clause that is contrastive with the remnant in the answer clause.

4. Conclusion

In this paper, we have argued that the sentence composed of the PAP *ung* 'yes' or *ani* 'no', and followed by one XP remnant (alias UAX) involves right dislocation. The PAP has been widely assumed as a particle, but it is, in fact, followed by the covert TP that has undergone TP ellipsis.

In this respect, this elided TP in addition to the PAP is necessary to the syntactic licensing of the ensuing RD-ed remnant and provides a right structural context for the right dislocation of the RD-ed remnant. We have also examined different types of UAX-RD and the parallelism between them and canonical RDs in terms of syntactic behaviors. Evidence for our proposal is drawn from a wide range of syntactic behaviors such as island sensitivity, the full host clause requirement, voice/verbal morphology matching, and specificational coordination. At the current stage of knowledge, we conclude that there is no difference between UAX and canonical RDs.

Name: Ui-Jong Shin
Affiliation: Dongguk University
E-mail address: shinuj1017@gmail.com

Name: Sunjoo Choi
Affiliation: Dongguk University
E-mail address: sunjoo3008@gmail.com

References

An, D. H. (2016). Extra deletion in fragments and its implications. *Journal of East Asian Linguistics*. On-line first final version available at the journal's website. DOI: 10.1007/s10831-016-9144-7.

Ahn, H. D. & S. Cho. (2015). Right dislocation vs. fragment: A reply to Ko (2014). *Studies in Generative Grammar* 25, pp. 427-446.

Ahn, H. D. & S. Cho. (2015). A uniform analysis of right dislocation: A reply to Ko (2016). *Language Research* 52:2, pp. 213-245.

Chung, D. H. (2009). Do Not Target a predicate: It's not a constituent. A paper presented at the 6th WAFL (Workshop on Altaic Formal Linguistics), Nagoya, Japan.

Chung, D. H. (2011). A Constituency-based explanation of syntactic restrictions on Korean predicates. *Linguistic Research* 28:1, pp. 393-407.

Chung, D. H. (2014). On the locus of the interpretable negative feature in a negative sentence. Language Research 50:2, pp. 461-486.

Chung, D. H. (2015). Some notes on Ko's(2014b, 2015) hybrid approach to the Korean RDC. *Studies in Generative Grammar* 25:3, pp. 735-754.

Griffith, J & A. Liptak. (2014). Contrast and island sensitivity in clausal ellipsis. *Syntax* 17:3, pp. 189-234.

Han, C. H. & J. B. Kim. (2004). Are there 'double relative clauses' in Korean? *Linguistic Inquiry* 35, pp. 315-337.

Hoji, H. (1990). Theories of Anaphora and Aspects of Japanese Syntax. Ms., University of Southern California, Los Angeles.

Holmberg, A. (2015). *The Syntax of Yes and No.* New York: Oxford University Press.

Im, H. B. & S. W. Chang (1995). *Kwuke Mwunpeplon* (Korean Grammatical Theory). Korea Correspondence University Press.

Kim S. Y & S. Kuno (2013). A note on Sluicing with implicit correlates, *Natural Language Semantics*, pp.315-332.

Ko, H. J. (2014). Right dislocation as specificational focus. Talk presented at *Workship on Korean-Japanese Linguitics at SNU*, Oct 17, 2014, Seoul National University.

Ko, H. J. (2015). Two ways to the right: A hybrid approach to right-dislocation in Korean. to appear in *Language Research* 51:1, pp. 3-40.

Ko, H. J. (2016). Gapless rght-dislocation: The role of overt correlates. *Language Research* 52:1, pp. 3-32.

Kuno, S. (1973) The structure of the Japanese language. Cambridge, MA: M.I.T. Press.

Kuno, S. (1978). *Danwa-no Bunpo* [Grammar of discourse], Tokyo: Taishukan Shoten.

Kramer, R, & K. Rawlins (2009). Polarity particles: an ellipsis account. In *Proceedings of the 39th Annual Meeting of the North East Linguistic Society*. Amherst, MA: Graduate Student Linguistic Association.

Lee, C. H. (2009). *Hankwuke Hwupochwung Kwumwun-uy Kwuco* (The Structure of Korean Afterthought Constructions). Emun-yenkwu 37, pp. 31-54.

Merchant, J. (2013). Voice and ellipsis. *Linguistic Inquiry* 44, pp. 77-108.

Park, M. K. (2015). Ways of replying to negative questions in English and Korean. *Korean Journal of English Language and Linguistics* 15, pp. 397-421.

Park, M. K. & U. J. Shin (2014). Embedded verbal complexes in Korean: Towards a unified approach to leftward movement, right dislocation and anaphora. *Linguistic Research* 31:3, pp. 569-590.

Park, M. K. & U. J. Shin, and Kang, J. M (2015). The syntax of multiple fragments in Korean. A paper presented at the 2015 17th SICOGG.

Simon, M. E. (1989). An analysis of the postposing construction in Japanese. Ph.D dissertation, University of Michigan.

Yoon, H. S. (2013). Implications of constraints on null constituents for analyses of the right dislocation construction. Department of English Colloquium, Dongguk University (6/12/13).

Yoon, J. M. (2011). Double relativization in Korean - An explanation based on the processing approach to island effects. *Korean Journal of Linguistics* 36:1, pp. 157-193

Cognate adverb construction in Moksha Mordvin

Sofia Nikiforova

This study deals with the semantics of the cognate adverb construction (CAC) in Moksha Mordvin. This construction is formed by a conjugated verb form and a deverbal adverb derived from the same root. I show that meanings of the Moksha CAC fall into two categories: intensification and so called 'accuracy of description'. Some typologically close phenomena are demonstrated. To conclude, a hypothesis is suggested that the CAC has a single 'underlying' meaning that can be described through use of the prototype notion.

1. Introduction

This study deals with the semantics of the cognate adverb construction in Moksha, a Finno-Ugric language (Uralic language family) spoken in Mordovia, Russia. This phenomenon has not been studied closely before, nor has it been compared to similar constructions in other languages. In this study the Moksha data is viewed from a typological perspective; I provide a thorough description of the construction's semantics and propose an analysis that can also be applied to other languages.

The cognate adverb construction (CAC) is formed by a conjugated verb form and a deverbal adverb ending in *-z'*, derived from the same root. The adverb itself has the meaning of an 'attendant circumstance' (König 1995), see (1).

(1) *c'ora-s'*   *mol'-i*   *mora-z'*
     boy-DEF.SG   walk-PRS.3SG   sing-CONV.ATD
     'The boy walks singing.'

But unlike these usual 'verb + adverb' combinations that have more or less compositional semantics, the construction in question possesses its own meaning, see (2).

(2) *c'ora-s'*   **ud-əz'**   **ud-i**   (lit.: the boy sleeps (how?) sleeping)
     boy-DEF.SG   sleep-CONV.ATD   sleep-NPST.3SG
     1. 'The boy sleeps soundly.'
     2. 'The boy is <u>sleeping</u> (and not just lying with his eyes shut).'

The data for my research was collected in 2015–2016 during field work with speakers of the Central dialect of the Moksha language in the villages of Lesnoje Cibajevo, Lesnoje Ardashevo, and Lesnyje Sijali in Mordovia (Russia).

This paper is structured as follows. In section 2 I present the basic syntactic properties of the Moksha construction. In section 3 I describe its 'surface' semantics. In section 4 I discuss the typological status of the CAC. In section 5 I draw conclusions and suggest a hypothesis regarding the construction's 'underlying' semantics.

2. Syntax

The standard word order in the cognate adverb construction is "adverb - verb". This is how the construction appears in a neutral context. The order can be reversed under specific pragmatic circumstances, namely when the semantic component that the cognate adverb brings about is in focus, see (3).

(3)   *c'ora-n'ɛ-s'*   ***las'k-i***   ***las'k-əz'***
       boy-DIM-DEF.SG run-PRS.3SG   run-CONV.ATD
       'The boy is running <u>fast</u>.'

Although it is not common, the cognate adverb and the verb can be separated from each other by various constituents, such as a noun phrase, as in (4). However, if the constituent is long and "heavy", like the noun phrase with a relative clause in (4), the separation is not acceptable.

(4)   a. *^{OK}son* ***mol'-əz'***   *t'ɛ*   *kud-t'i*   ***mol'-i***
          he   walk-CONV.ATD that house-DEF.DAT   walk-PRS.3SG
          'He is <u>walking</u> {not running} to that house.'

       b. **son* ***mol'-əz'***   *kud-t'i*   *kona*   *ašč-i*   *ičkəz'ən'*
          he   walk-CONV.ATD   house-DEF.DAT   that   be.situated-PRS.3SG   faraway
          pandə-n'ɛ-t'   *esə*   ***mol'-i***
          hill-DIM-DEF.GEN   in.IN   walk-PRS.3SG
          Expected meaning: 'He is <u>walking</u> {not running} to the house that stands on the faraway hill.'

The cognate adverb can also receive the additive particles *-gə* and *-əvək*, which cannot be attached to the adverb ending in *-z'* outside the construction (as stated in Kozlov 2015).

(5) ***mora-z'-gə***   ***mora-k****!*
    sing-CONV.ATD-ADD   sing-IMP.SG
    '{"she is telling a story, and somebody tells her: do not just tell it, "} <u>Sing</u> as well!'

There is one more difference between a simple adverb ending in *-z'* and a cognate adverb in the CAC. An adverb with the meaning of 'attendant circumstance' cannot be derived from achievements (in the sense of Vendler 1957), while a cognate adverb in the construction can, see (6), with the achievement verb *komət'əms* 'to jump'.

(6)   a. **son* ***komət'-əz'***   *koṛn'-i*
          he   jump-CONV.ATD talk.IPFV-PRS.3SG
          Expected meaning: 'He is talking (while) jumping/having jumped.'

       b. *son* ***komət'-əz'***   ***koməc'***
          he   jump-CONV.ATD jump.PST.3SG
          'He <u>jumped</u> {didn't just take a step}.'

In the Erzya language, which is closely related to Moksha, the suffix *-z'* forms an adverb that denotes not only 'attendant circumstances' but also resultative states. Such adverbs can easily be derived from achievements. This suggests that the constraint on the derivation of the Moksha *z'*-adverb is a relatively recent development and it is possible that the cognate adverb construction preserved the otherwise lost usage.

3.   Semantics

Meanings of the Moksha cognate adverb construction fall into two main categories: intensification and accuracy of description.

3.1. Intensification

Different groups of verbs receive different kinds of intensification. Many verbs can be intensified with respect to the most prominent parameter in their semantic class. Here I will list those different types (subclasses) of intensification that I have been able to distinguish.

- 'quickly'

Almost all motion verbs acquire this meaning, with one of the exceptions being the verb *šaštəms* 'to crawl', which describes prototypically slow motion.

(7) *c'ora-s'*   **las'k-əz'**   **las'k-i**   *ki-t'*   *ezga*
      boy-DEF.SG run-CONV.ATD   run-PRS.3SG   road-DEF.GEN   in.PROL
      'The boy is running fast along the road.'

(8) *narmət'-t'n'ə*   **l'ij-əz'**   **l'ij-s't'**   *oš-t'*   *šir'-i*
      bird-DEF.PL   fly-CONV.ATD   fly-PST.3PL town-DEF.GEN   side-LAT
      'Birds have flown quickly on the way to the town.'

(9) *vec'*   **šud'-əz'**   **šud'-i**   *kran-t'*   *ezdə*
      water.DEF.SG   pour-CONV.ATD   pour-PRS.3SG   faucet-DEF.GEN   in.ABL
      'The water is pouring down quickly from the faucet.'

- 'slowly'

(10) *id-n'ɛ-s'*   **šašt-əz'**   **šašt-i**   *mastər-ga*
      child-DIM-DEF.SG   crawl-CONV.ATD   crawl-PRS.3SG   floor-PROL
      'The child is crawling slowly across the floor.'

Neutral motion verbs *mol'əms* 'to go, to walk' and *jakams* 'to walk' that denote neither fast nor slow movement are not intensified in the CAC.

- 'loudly'

Almost all sound verbs acquire this meaning, with one of the exceptions being the verb *toškams* 'whisper' that describes a prototypically quiet sound.

(11) *al'ɛ-s'*   *ud-i*   *i*   **kərna-z'**   **kərna-j**
     man-DEF.SG   sleep-PRS.3SG   and   snore-CONV.ATD   snore-PRS.3SG
     'The man is sleeping and snoring loudly.'

(12) *katə-n'ɛ-s'*   *pen'akud-t'*   *lanksə* **mərna-z'**   **mərna-j**
     cat-DIM-DEF.SG   heater-DEF.GEN   on.IN   purr-CONV.ATD   purr-PRS.3SG
     'The cat on a heater is purring loudly.'

(13) *id-n'ɛ-s'*   *toka-v-s'*   *i*   **avar'd'-əz'**   **avar'd'-i/**
     child-DIM-DEF.SG hit-PASS-PST.3SG and   cry-CONV.ATD   cry-PRS.3SG
     ivad'-əz'   **ivad'-i**   *(što…)*
     shout-CONV.ATD   shout-PRS.3SG   that
     'The child hurt himself and is crying loudly/shouting loudly (that…)'

(14) *mastər-s'*   *t'ɛ* *kud-t'*   *esə*   **čatərd-əz'**   **čatərd-i**.
     floor-DEF.SG   this house-DEF.GEN in.IN   squeak-CONV.ATD   squeak-PRS.3SG
     mon *af*   *matəd-əv-an*   *kədə* *kucə*   *ki-vək*   *jaka-j*
     I   NEG   fall_asleep-PASS-PRS.1SG   if   house.IN who-ADD   walk-PRS.3SG
     'The floor in this house squeaks very loudly. I cannot fall asleep if someone is walking in the house.'

- 'softly'

(15) *mes*   *ton*   **toška-z'**   **toška-t**,   *koṟta-k*   *normal'naj-stə!*
     why   you   whisper-CONV.ATD   whisper-PRS.2SG   talk-IMP.SG normal-EL
     'Why are you whispering so softly, talk normally!'

Not all verbs that denote a situation accompanied by some kind of sound emission can be intensified. For example, speech verbs (except for *toškams* 'to whisper' and *ivad'əms* 'to shout') cannot be intensified.

- intensification of the resultative component; 'fully, completely'

In the case of verbs with an incremental theme (Dowty 1991), the CAC indicates that the whole theme is affected by the event. If the verb is telic, then it also means that the action has completed; this semantic affect does not appear in the case of atelic verbs.

(16) *mon'*   *sas'ed-ən'ə-n'*   *esə*   *ul'-s'*   *požar.*
     I.GEN   neighbour-1SG.POSS.PL-GEN   in.IN   be-PST.3SG fire
     kuc   **pal-əz'**   **pal-s'**   *kaftə*   *čast-t*
     house.DEF.SG burn-CONV.ATD   burn-PST.3SG   two   hour-PL
     'There was a fire in my neighbour's house. The whole house has been {fiercely} burning for two hours.'

(17) *kenkš-s'*   *varma-t'*   *ezdə*   **panž-əz'**   **panž-əv-s'**
     door-DEF.SG   wind-DEF.GEN   in.ABL   open-CONV.ATD   open-PASS-PST.3SG
     'Because of the wind the door {quickly} opened wide.'

This meaning of the CAC is usually followed by some kind of implicature. This implicature (generalized implicature in the terms of Grice (1975)) can be removed by a specific context, but in a neutral context it is practically obligatory, unlike in sentences containing the word *marn'ək* 'fully, completely' instead of the CAC.

The implicature can be the semantic component 'quickly' or, by contrast, 'thoroughly', 'gradually'. The kind of implicature that appears in any given case depends on the semantics of the verb and, even more so, on the pragmatic context and does not depend on any grammatical properties of the verb.

(18) a. *sa-s'   tunda-s'   i   lov-s'   **sola-z'**   **sola-s'***
         come-PST.3SG   spring-DEF.SG and snow-DEF.SG melt-CONV.ATD melt-PST.3SG
         'Spring has come, and all the snow has {immediately} melted.'

      b. *t'ɛn'i   tunda   i   lov-s'   **sola-z'**   **sola-j***
         now   spring and snow-DEF.SG melt-CONV.ATD   melt-PRS.3SG
         'It is spring, and the snow is melting fast everywhere.'

(19) a. *vandi   ul'-i   praz'n'ik. mastər-s'   **šta-z'**   **šta-f** /*
         tomorrow be-PRS.3SG celebration floor-DEF.SG   wash-CONV.ATD   wash-PTCP.RES
         narda-z'   **narda-f***
         wipe-CONV.ATD   wipe-PTCP.RES
         'There will be a celebration tomorrow. The whole floor is completely washed/wiped.'

      b. *mon **šta-z'**   **šta-sa**   mastər-t'*
         I   wash-CONV.ATD wash-PRS.1SG.S.3SG.O   floor-DEF.GEN
         'I am washing the floor thoroughly {every inch}.'

(20) *son   c'ebɛr' rabota t'ij-s',   son   **aṛt-əz'**   **aṛt-əz'ə***
      he   good   job   do-PST.3SG he   paint-CONV.ATD paint-PST.3SG.S.3SG.O
      zabor-t',   s'embə doska-n'a-tn'ən   aṛt-əzn'ə
      fence-DEF.GEN   all   доска-DIM-DEF.PL.GEN   paint-PST.3SG.S.3PL.O
      'He has done a good job: he has thoroughly painted the fence, all the planks.'

(21) *c'ora-s'   **kala-ft-əz'**   **kala-ft-əz'ə**   tarad-t'*
      boy-DEF.SG   break-CAUS-CONV.ATD   break-CAUS-PST.3SG.S.3SG.O branch-DEF.GEN
      'The boy has gradually broken the whole branch {first in two pieces, then in four, and so on}.'

- intensification of frequency; 'all the time'

In Moksha, habitual semantics is expressed by special verbal suffixes. If a verb with one of these suffixes appears in the CAC, then it acquires the meaning 'all the time (not just sometimes)'. It is usually accompanied by a negative connotation (cf. English use of the present participle in cases such as 'he is always doing this!').

(22) *son   **suva-z'**   **suv-s'-i**   mon'   kud-əzə-n   apak*
      he   enter-CONV.ATD enter-FREQ-PRS.3SG I.OBL   house-ILL-1SG.POSS NEG.NFIN
      kiz'əft
      ask.CN.NFIN
      'He is always entering my house without asking.'

(23) *t'a-sə   af   c'ebɛr' vastə, **naksad-əz'**   **naksačn'-i**   modamar'-s'***
      this-IN NEG good   place   rot-CONV.ATD   rot.FREQ-PRS.3SG potato-DEF.SG
      'It is a bad place, potatoes always rot here.'

(24) *t'ɛ*   *kruška-s'*   **pra-z'**   **pra-šənc'**   *polka-t'*   *lank-stə*
     this mug-DEF.SG   fall-CONV.ATD   fall-FREQ.PST.3SG   shelf-DEF.GEN   on-EL
     məz'ar-s   *iz'-in'ə*   *put-ə*   *šir'-i*
     when-ILL   NEG.PST-PST.1SG.S.3SG.O   put-CN   side-LAT
     'The mug kept falling down from the shelf until I put it away.'

(25) *son*   **ašč-əz'**   **ašč-əkšn'-i**   *kucə*
     he   stay-CONV.ATD   stay-FREQ-PRS.3SG   house.IN
     'He is at home all the time.'

- intensification of extent; 'to a great extent'

All verbs in this class denote a state that can be present to a greater or lesser extent. The use of the construction indicates that the degree to which the state is present is quite high. This class includes many verbs of different semantic classes; here I list only three examples.

(26) *mon*   **jora-z'**   **joran**   *ud-əm-s*
     I   want-CONV.ATD   want.PRS.1SG   sleep-INF-ILL
     'I want to sleep very much.'

(27) *son*   **kel'g-əz'**   **kel'k-si**   *es'*   *ava-nc*
     he   love-CONV.ATD   love-PRS.3SG.S.3SG.O   own   wife-3SG.POSS.SG.GEN
     'He loves his wife very much.'

(28) *urok-s'*   *c'ebɛ r'-əl',*   *učen'ik-n'ə*   *s'embə*   **kuɬcənd-əz'**   **kuɬcənc't'**
     lesson-DEF.SG good-PQP   student-DEF.PL   all   listen-CONV.ATD listen.PST.3PL
     'The lesson was good, all students listened very attentively.'

- intensification of quality; 'very well'

Finally, there are verbs for which the use of construction means that the action is carried out well, or rather, like it *should* be carried out (according to the speaker's worldview).

(29) *mərapr'ijat'ijə-s'*   **jota-z'**   **jota-s'**
     event-DEF.SG   walk-CONV.ATD   walk-PST.3SG
     'The event went very well.'

(30) *garmon'ija-s'*   **naɬk-əz'**   **naɬk-i**
     accordion-DEF.SG   play-CONV.ATD   play-PRS.3SG
     'The accordion is playing very well.'

3.2. Accuracy of description

Not all verbs become intensified when they appear in a CAC; however, given a context of implicit or explicit contrast, all Moksha verbs in this construction can receive the 'accuracy of description' meaning.

   Describing a situation, speakers always choose the expression that they find the most suitable for achieving their communicative goals. The second meaning of the Moksha construction is to indicate the speaker's certainty that the choice of verb has been as good as possible, that the use of any other verb would lead to a misunderstanding between the speaker and the listener. This meaning is illustrated in examples (31)–(35).

(31) *ava-s'*   ***šač-ft-əz'***   ***šač-ft-əz'ə***   *it'-t'*
      woman-DEF.SG be_born-CAUS-CONV.ATD   be_born-CAUS-PST.3SG.S.3SG.O   child-DEF.GEN
      1. 'The woman <u>bore</u> a child {she did not adopt him}.'
      2. 'The woman <u>bore</u> a child herself {she did not have a caesarean section}.'

(32) *son'*   *kulə-ft-əz'?*   –   *ajaf, urma-sə*   ***kulə-z'***   *kulə-s'*
      he.OBL   die-CAUS-CONV.ATD   no   illness-IN   die-CONV.ATD die-PST.3SG
      'Has he been killed? – No, he <u>died</u> of an illness.'

(33) *son* ***s'ormad-əz'***   ***s'ormad-əz'ə***   *s'orma-t'*
      he   write-CONV.ATD write-PST.3SG.S.3SG.O letter-DEF.GEN
      'He <u>wrote</u> the letter {he didn't type it}.'

(34) *mes   ton   mora   azə-nd-at?*   ***mora-z'***   ***mora-k***!
      why   you   song   tell-FREQ-PRS.2SG   sing-CONV.ATD   sing-IMP.SG
      'Why are you telling me a song? Sing it!'

(35) *son* ***kand-əz'***   ***kand-əz'ə***   *id-t'*
      he   carry-CONV.ATD   carry-PST.3SG.S.3SG.O   child-DEF.GEN
      'He <u>carried</u> the child {he didn't push him in a pram}.'

This meaning of the CAC is often found in metaphors and exaggerations. The use of metaphors suggests that the speaker considers it to be the most suitable to describe this situation using an expression that is usually used in a different context. And that is exactly what is conveyed through the use of this construction. The contrast in this case is with other, non-metaphorical ways to describe a situation.

(36) ***l'ij-əz'***   ***l'ij-s'***   *od   pingə-z'ə*
      fly-CONV.ATD   fly-PST.3SG   young   time-1SG.POSS.SG
      'My young years have really <u>flown by</u> {not just gone by}.'

(37) *mər'd'ə-z'ə*   ***ped'-əz'***   ***pec'***   *futbol'naj matč-t'i*
      husband-1SG.POSS.SG stick-CONV.ATD   stick.PST.3SG football   match-DEF.DAT
      'My husband has really <u>stuck</u> to a football match {he is so hooked}.'

(38) *varma-s'*   ***ker'ə-z'***   ***ker'ə-saman'***
      wind-DEF.SG cut-CONV.ATD   cut-PRS.3SG.S.1SG.O
      'The wind is just <u>cutting</u> me {it is that strong}.'

(39) *son nacianal'naj-t'*   *inksə*   ***pal-əz'***   ***pal-i***
      he   national-DEF.GEN   because_of   burn-CONV.ATD   burn-PRS.3SG
      'He is <u>burning</u> because of the nationality issue {because he cares so much}.'

(40) *fkɛ   pink-t'*   ***narda-z'***   ***narda-v-s't'***   *s'embə val̡-n'ə*
      one   time-DEF.GEN   wipe-CONV.ATD   wipe-PASS-PST.3PL   all   word-DEF.PL
      'At once all the words have been <u>wiped out</u> {= forgotten}.'

4. Typological status

In this section I will address the issue of the typological status of the Moksha construction.

In order to compare it to similar constructions in other languages, first I have to decide which constructions can be regarded as similar. From a semantic point of view, it is perhaps closest to some cases of reduplication. Intensification is always listed among meanings that reduplication can convey (see, among others, Moravcsik 1978; Regier 1994; Hurch 2005). The 'accuracy of description' meaning is rarer, but not completely unattested (cf. Ghomeshi et al.(2004) on so called contrastive focus reduplication in English in sentences such as 'it is not a tuna salad, it is a SALAD-SALAD').

Of course, the fact that the Moksha CAC is formed by root repetition is also an argument for it to be considered a reduplication. However, strictly speaking, the form of the construction is precisely why it cannot be considered one. Traditionally, reduplication is thought to be a morphological process that results in conveying some lexical or grammatical meaning through segment repetition. This is how reduplication was treated in, for example, Marantz (1982); Haspelmath (2002). Only a few researchers allowed cases like this one, when repetition happens on a syntactic rather than morphological level, to be called 'reduplication'. One of these few is Maas (2005) who discusses an Arabic 'syntactic reduplication construction'. What he describes is actually a phenomenon that has been known to linguists for a long time as *figura etymologica,* that is verb+noun constructions such as 'to die a quiet death', 'to smile a happy smile'. In modern linguistics constructions of this type are often referred to as Cognate Object Constructions (see, e.g., Huddleston and Pullum 2002). Their meaning, however, is quite far from the meaning of the Moksha Cognate Adverb Construction (which, of course, has been named precisely after the Cognate Object Construction).

So, it seems that no typological label can be applied to the Moksha CAC without cause for disagreement. Of course, it does not mean that it is typologically unique; it only points out the need to ignore the strict boundaries some researchers may have created. In this study I am more interested in semantics, and that is why in the remainder of this chapter I will focus more on constructions that have similar semantics to those of our Moksha construction; formal similarity (root repetition) is necessary but not sufficient for being taken into account.

4.1. Uralic languages

Almost identical constructions can be found in languages closely related to Moksha Mordvin: Erzya Mordvin and Shoksha Mordvin. They also consist of a verb and an 'attendant circumstance' adverb with the same root. We can see from examples below that both meanings (intensification and 'accuracy of description') present themselves in Erzya Mordvin and Shoksha Mordvin.

(41)   Erzya Mordvin (Cygankin 1980:359)
        *ob'ed.   či-s'   **p'id'-əz'**   **p'id'-i**
        dinner   sun-DEF.SG bake-CONV.ATD   bake-NPST.3SG
        'Dinner time. The sun is very hot {lit. bakes baking}.'

(42)   Erzya Mordvin (Kolyadenkov, Zavodova 1962:325)
        sev-əz'   **sev-it'**   *es-ən'*
        eat-CONV.ATD   eat-NPST.3PL   IN-1SG
        'They are just <u>eating</u> me alive {=scolding me harshly}.'

(43)   Shoksha Mordvin (Popova 2007:12)
        *son   **l'ift'e-z'**   **l'ift'e**
        he   fly-CONV.ATD   fly.3SG
        'He is flying very quickly.'

(44)   Shoksha Mordvin (Popova 2007:13)
        *son   **soka-z'**   **sok'e**   paks'a-t'e-nde*
        he   plow-CONV.ATD   plow.3SG   field-DEF-PART
        'He is plowing the field very well.'

(45)   Shoksha Mordvin (Popova 2007:13)
        *son   at   **mol'i-z'**   **mol'e**   a   **ardu-z'**   **arde***
        he   NEG   walk-CONV.ATD   walk.3SG   but   ride-CONV.ATD   ride.3SG
        'He is not <u>walking</u> but <u>riding</u>.'

In Hungarian there is a syntactic reduplication construction formed by a deverbal adverb derived from the root by the suffix *-va/-ve*, followed by the conjugated form of the same verb, see (46). It is used to express intensification.

(46)   Hungarian (Brdar et al. 2014:5)
        *hát   **kér-ve**   **kér-ünk**   titek-et*
        well   ask-ADV   ask-1PL   2PL-ACC
        'We beg you <u>very much</u>.'

<div align="center">

4.2. Semitic languages

</div>

So-called tautological infinitive constructions are found in many, if not all, Semitic languages (Biblical Hebrew, Old Aramaic, Syriac, Ugaritic, Akkadian, Classical Arabic, Maltese and others). These constructions are formed by a conjugated verb and a special verbal form usually called 'the infinitive absolute' ('masdar' in Arabic linguistics).

  The tautological infinitive construction is most commonly used to express the certainty of an event (as shown in, for example, Allen 2010; Kim 2009; Pereltsvaig 2001; Goldenberg 1971; Callaham 2010). It should be noted that it is not the same meaning as the Moksha 'accuracy of description' - the Moksha meaning specifies <u>what</u> situation has occurred, while the Semitic meaning states that the situation <u>has</u> (or will have) occurred. This 'assertion intensification' meaning is illustrated in examples (47)–(52).

(47)   Biblical Hebrew (Prokopenko 2013:157)
         *we'averaham   **hayoh**   **yihyeh**   legoy   gadol   we'atsum*
         and=Abraham   be.INF.ABS   be.FUT.3M   nation   big   and=strong
         'And Abraham will <u>definitely</u> become a mighty and strong nation.'

(48)   Biblical Hebrew (Anstey 2003)
        homleach   *lo*   ***humlachat***
         rob:PASS.CAUS.INF.ABS   NEG   rob:PASS.CAUS.2M.SG
         'You were <u>definitely</u> not robbed.'

(49)   Classical Arabic (Yasin 2014:331)
        đarab-t-u   *Zaid-an*   ***đarb-an***
         hit:PERF-1S-IND   Zaid-ACC   hitting-ACC
         'I <u>did</u> hit Zaid.'

(50)   Ugaritic (Waltke et al. 1990:584)[1]
        ǵmʾu   ǵmʾit
         'You (fem.) are <u>certainly</u> thirsty.'

---

[1] Unfortunately, examples are not glossed in (Waltke et al. 1990), (Kogan, Lyozov 2009), (Kogan 2009) and (Lyozov 2009).

(51) Akkadian (Kogan, Lyozov 2009:168)
    tabālum   tatbal
    'You <u>really did</u> take.'

(52) Syriac (Lyozov 2009b:600)
    balḥod̲   mhaymånu   haymen
    just   believe.INF   believe.IMP
    'Just <u>do believe</u>!'

The tautological infinitive construction is also used to show the completeness of an occurrence, see (53)–(55).

(53) Ugaritic (Kogan 2009: 220)
    ḫtʔi   [ḫatāʔ]   nḫtʔu
    'They have been completely destroyed.'

(54) Biblical Hebrew (Gesenius et al. 1922:313)
    wehoreš   lo   horišo
    and=drive_out.INF.ABS NEG drive_out.PST.3PL.O
    'And did not utterly drive them out.'

(55) Classical Arabic (Maas 2005:410)
    ʃabiʕ-u: ʃabʕ-an
    become.satisfied:PF-3.PL   satiety-ACC.SG.-DEF
    'They became completely satisfied.'

The construction can intensify one of the verb's semantic components in various other ways, see (56)–(60).

(56) Old Aramaic (Lyozov 2009a:444)
    hwm ʔthmw
    ʿThey became very puzzled.'

(57) Biblical Hebrew (Gesenius et al. 1922:313)
    šim'u   šamoa
    hearken.INF.ABS hearken.IMP
    'Hearken ye attentively.'

(58) Maltese (Maas 2005: 416)
    a. *ja-ħbat   ħbit*
      3.S.M-strike:IPF   striking
      'He has a lot of accidents.'

    b. *ħabat   ħabta*
      strike:PF.3.S.M   striking
      'He had a big accident.'

(59) Classical Arabic (Maas 2005:410)
    ʔiqtatal-na: qita:l-an
    kill:REFL:PF-PL   killing-ACC.SG-DEF
    'We fought hard.'

(60)   Classical Arabic (Maas 2005:410)
         ja-rin:-u   r-rani:n-a
         3-ring:IPF-P DEF-ringing-ACC
         'He rings the bell intensively.'

4.3. Lango (Nilo-Saharan)

In Lango (Western Nilotic < Nilotic < Nilo-Saharan) there is a construction formed by a verb and a gerund with the same verbal root. The Lango construction seems to have both the Semitic 'assertion intensification' meaning and the Moksha 'accuracy of description' meaning, see examples (61)–(62).

(61)   Lango (Noonan 1992:175)
         àbínô   àbínə̂   àwó'ró
         1SG.come.PERF   come.GER   yesterday
         'I <u>did</u> come yesterday.'

(62)   Lango (Noonan 1992:175)
         *ɲákô   **òmyɛ̀lò   àmyɛ́lâ***
         girl   3SG.dance.PERF   dance.GER
         'The girl <u>just danced</u>.'

4.4. Russian (Slavic)

In Russian (Slavic) there are two different verbal constructions with root repetition.

   The first consists of a conjugated verb immediately preceded by an adverbial formation with the suffix *-mya* or *-om/-em*. In Shvedova (1960); Shvedova (1980); Mrazek (1964) this construction is described to have an effect of 'intensifying the verbal semantics'. Some examples are given below.

(63)   Russian
         medved **rev-mya   rev-et**
         bear   roar-ADV   roar-PRS.3SG
         'The bear roars loudly.'

(64)   Russian
         oxotnik **beg-om   po-bež-al**
         hunter   run-ADV   INCH-run-PST.3SG.M
         'The hunter started running fast.'

(65)   Russian
         ona goryu-et,   **plač-em   plač-et**
         she   grieve-PRS.3SG   cry-ADV   cry-PRS.3SG
         'She is grieving, crying hard.'

Apart from intensification, it can also convey the 'accuracy of description' meaning, although it is much rarer in contemporary speech.

(66)   Russian (Goncharov 1842:14)
         ona čašš'e   by-va-la   *u nego: obeda-la,*
         she   often.CMPR   be-FREQ-PST.3SG.F   at he.GEN   have_dinner-PST.3SG.F
         zavtrakal-a   —   slovom,   kak govor-yat,   ***živ-mya   ži-la***
         have_breakfast-PST.3SG.F word-INSTR   as   say-PRS.3PL   live-ADV   live-PST.3SG.F

'She was at his place more often: had dinner [there], had breakfast - in one word, as they say, <u>really lived</u> [there].'

The second construction also contains root repetition - the verb is repeated in exactly the same form and the two verbs are separated by the conjunction *tak* 'so'. In Kopotev (2009) this construction is called 'the phraseme of the utmost quality', meaning that it serves to indicate 'a statement, accepting of an identified, or offered object, but in its best, fullest, most qualitative manifestation' (cited in Kopotev (2009) from Velichko 1996). It is usually perceived as, again, intensification, see (67)–(68).

(67)   Russian
       *on   togo   mužik-a   **udari-l**   **tak**   **udari-l**
       he   that   man-ACC   hit-PST.3SG.M   so   hit-PST.3SG.M
       'He hit that man very hard.'

(68)   Russian
       ***gulya-t'**   **tak**   **gulya-t'**!*
       party-INF   so   party-INF
       'Let's party hard!'

5.   *Conclusions and hypothesis*

Having fully described the semantics of the Moksha cognate adverb construction from a typological point of view, in this section I will try to bring the facts together to form a general hypothesis regarding the underlying semantics of the construction.

My hypothesis is that the two meanings of the Moksha CAC are in fact two manifestations of a single meaning. This meaning is defined as follows: the construction narrows down the verb denotation leaving only its nuclear part, its most prototypical representatives.

In the semantics of some verbs there is a parameter that is crucial for distinguishing between prototypical and non-prototypical situations denoted by the verb. It is, for example, the speed parameter for the verb 'to run': fast running is prototypical and slow running is not. When a speaker uses a verb like 'to run' in a CAC, only prototypical (i.e. fast) running is meant. Therefore, it brings about the intensification effect.

If there is no such parameter in the verb's semantics, there is no intensification effect. However, the use of a CAC still has the same meaning, namely that the situation the speaker is describing is prototypical of the verb in the construction. This creates an implicature: 'if the situation is prototypical of the verb V, there can be no mistake in choosing V for its description' and thus leads to the 'accuracy of description' meaning.

There is another difference found in using cognate adverb constructions with these two types of verbs. Let us first consider those verbs that have some parameter, the value of which is rising while moving from the periphery of the verb denotation to its nuclear part. These verbs (e.g. 'to run', 'to shout', 'to love') can be used in a CAC without a context of contrast. However, contrast is necessary in the case of the other type of verbs, that is the ones that do not have such a parameter. Verbs such as 'to kill' or 'to buy' require a context of contrast (although it may be implicit) to be used in a CAC. The reason for this is following general rules of discourse organization, namely Grice's Maxim of Quantity, 'be as informative as required' (Grice 1975). If a verb's semantics includes a salient enough descriptive characteristic that separates good, prototypical denotation representatives from not so prototypical ones, then the assertion of prototypicality provided by the CAC is informative on its own. But if it does not, then this assertion is not only meaningless, it is also uninterpretable. Here is when contrast comes to play: based on, first, the listener's assumed doubt that the chosen verb describes the situation best of all, second, the assertion of prototypicality,

and, third, Grice's Maxim of Relation ('be relevant'), the following implicature is created: 'I could not be wrong in choosing the verb V'.

If a sentence with a CAC is negated, then what is really negated is again the assertion of prototypicality. It can point to either a non-prototypical realization of the situation or an absence of this situation altogether, and so the ambiguity occurs, as we can observe in example (69).

(69) *t'ɛ   af   vid'ə što   son   ezdə-nzə   pel'-əz'   pel'-i.*
      this NEG true that he   in.ABL-3SG.POSS be_afraid-CONV.ATD be_afraid-PRS.3SG
      (pel'-i   no   af   pɛk)
      be_afraid-PRS.3SG but   NEG   very
      1. 'It is not true that he is very afraid of him. {He is afraid, but not very much.}'
      2. 'He is not afraid of him.'

Data from different Uralic, Semitic and other languages mentioned above can be analyzed in a similar way, which is a matter for further research.

The proposed analysis puts Moksha CAC next to other cases of lexical and syntactic reduplication that are sometimes described within the prototype framework (cf. Gilyarova (2010) on noun reduplication in Russian).

Abbreviations

1	First Person		INF	Infinitive
2	Second Person		INSTR	Instrumental Case
3	Third Person		IPF	Imperfective
ABL	Ablative Case		FREQ	Frequentative
ABS	Absolute Form		LAT	Lative Case
ACC	Accusative Case		M	Masculine Gender
ADD	Additive Particle		NEG	Negation
ADV	Adverbalizer		NFIN	Non-finiteness
CAUS	Causative		NPST	Non-past Tense
CMPR	Comparative		O	Object
CN	Connegative		OBL	Oblique Case
CONV.ATD	Attendant Circumstance		PASS	Passive
DAT	Dative Case		PART	Partitive Case
DEF	Definiteness		PERF	Perfective
DIM	Diminutive		PF	Perfective
EL	Elative Case		PL	Plural
F	Feminine Gender		POSS	Possessive
FUT	Future Tense		PQP	Pluperfect
GEN	Genitive Case		PROL	Prolative Case
GER	Gerund		PRS	Present Tense
ILL	Illative Case		PST	Past Tense
IMP	Imperative		PTCP	Participle
IN	Inessive Case		RES	Resultative
INCH	Inchoative		S	Subject
IND	Indicative		SG	Singular

Sofia Nikiforova
HSE, Moscow
son.nik@mail.ru

References

Allen, D. L. (2010). *Hebrews: An Exegetical and Theological Exposition of Holy Scripture*. B&H Publishing Group.

Anstey, M. (2003). Post at Lingtyp. http://listserv.linguistlist.org/pipermail/lingtyp/2003-December/001360.html

Brdar, M. & R. Brdar-Szabó (2014). Syntactic reduplicative constructions in Hungarian (and elsewhere): Categorization, topicalization and concessivity rolled into one. Rundblad, G., A. Tytus, O. Knapton & C. Tang (eds.), *Selected Papers from the 4th UK Cognitive Linguistics Conference* 2, UK Cognitive Linguistics Assn., London, pp. 36-51.

Callaham, S. N. (2010). *Modality and the Biblical Hebrew Infinitive Absolute*. Harrassowitz Verlag, Wiesbaden.

Cygankin D. V. (ed.) (1980). *Grammatika mordovskih yazykov: fonetika, grafika, orfografiya, morfologiya*. Mordovskij gosudarstvennyj universitet, Saransk.

Dowty, D. (1991). Thematic Proto-Roles and Argument Selection. *Language* 67:3, pp. 547-619.

Gesenius, H. F. W., E. Kautzsch, A. E. Cowley, J. Euting & M. Lidzbarski. (1922). *Gesenius' Hebrew Grammar.* Oxford University Press, Oxford.

Ghomeshi, J., R. Jackendoff, N. Rosen & K. Russell (2004). Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory* 22, pp. 307-357.

Gilyarova, K. A. (2010). Takaya devochka-devochka. Semantika reduplikacii sushhestvitel'nyh v russkoj razgovornoj rechi i yazyke interneta. *Computational Linguistics and Intellectual Technologies Papers from the Annual International Conference 'Dialogue'* 16:9, RGGU, Moscow, pp. 90-96.

Goldenberg, G. (1971). Tautological infinitive. *Israel Oriental Studies* 1, pp. 36-85.

Goncharov, I. A. (1842). *Ivan Savich Podzhabrin*. Saint Petersburg.

Grice, H. P. (1975). Logic and conversation. Cole, P. & J. L. Morgan (eds.), *Syntax and Semantics: Speech Acts* 3, Academic Press, New York, pp. 22-40.

Haspelmath, M. (2002). *Understanding Morphology*. Oxford University Press, New York.

Huddleston, R. & G. K. Pullum (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press, Cambridge.

Hurch, B. (ed.) (2005). Studies on reduplication. *Empirical approaches to language typology* 28, Mouton de Gruyter.

Kim, Y. (2009). The function of the tautological infinitive in classical Biblical Hebrew. *Harvard Semitic Studies* 60. Eisenbrauns.

Kogan, L. E. (2009). Ugaritskij yazyk. *Yazyki mira: Semitskie yazyki. I. Akkadskij yazyk. Severozapadnosemitskie yazyki*. Academia, Moscow.

Kogan L. E. & S. V. Lyozov. (2009). Akkadskij yazyk. *Yazyki mira: Semitskie yazyki. I. Akkadskij yazyk. Severozapadnosemitskie yazyki*. Academia, Moscow.

Kolyadenkov M. N. & R. A. Zavodova (1962). *Grammatika mordovskih (mokshanskogo i erzyanskogo) yazykov: v 2 t. T. I. Fonetika i morfologiya*. Mordovskoe knizhnoe izdatelstvo, Saransk.

Kopotev, M. (2009). *Principles of syntactic idiomatization*. University of Helsinki.

Kozlov, A. A. (2015). *Zametki o fokusnyh chasticah*. Ekspedicionnyj doklad.

Lyozov S. V. (2009a). Aramejskie yazyki. *Yazyki mira: Semitskie yazyki. I. Akkadskij yazyk. Severozapadnosemitskie yazyki*. Academia, Moscow.

Lyozov S. V. (2009b). Klassicheskij sirijskij yazyk. *Yazyki mira: Semitskie yazyki. I. Akkadskij yazyk. Severozapadnosemitskie yazyki*. Academia, Moscow.

Noonan, M. (1992). *A grammar of Lango*. Walter de Gruyter, Berlin/New York.

Maas, U. (2005). Syntactic reduplication in Arabic. Hurch, B. (ed.), *Studies on reduplication. Empirical approaches to language typology* 28. Mouton de Gruyter, pp. 395-429.

Marantz, A. (1982). Re Reduplication. *Linguistic Inquiry* 13, pp. 435-482.

Moravcsik, E. (1978). Reduplicative Constructions. Greenberg, J. H. (ed.), *Universals of Human Language: Word Structure* 3, Stanford University Press, Stanford, pp. 297-334.

Mrazek, R. (1964). Designatory intensifikacii dejstviya v russkom yazyke. *Sborník prací Filozofické fakulty brněnské university*. Řada jazykovědná, pp. 169-179.

Pereltsvaig, A. (2001). Cognate objects in Modern and Biblical Hebrew. Ouhalla J. & U. Shlonsky (eds.), *Themes and Issues in Arabic and Hebrew*. Kluwer Academic Publishers, the Netherlands, pp. 1-33.

Popova, D. (2007). *Serialnye konstrukcii v shokshinskom dialekte mordovskogo yazyka: Ekspedicionnyj otchyot*. Moscow.

Prokopenko A. (2013). Obzor sintaksisa drevneevrejskogo yazyka. www.propovedi.ru/resource/hebrew-syntax-2013/download-pdf

Regier, T. (1994). *A preliminary study of the semantics of reduplication*. [Technical Report TR-94-019]. International Computer Science Institute, Berkeley, California.

Shvedova, N. Yu. (1960). *Ocherki po sintaksisu russkoj razgovornoj rechi*. Izdatelstvo Akademii nauk SSSR, Moscow.

Shvedova, N. Yu. (1980). *Russkaya grammatika. V 2-kh tomakh*. Izdatelstvo Nauka.

Velichko, A. V. (1996). *Sintaksicheskaja frazeologija dlja russkix i inostrancev*. Moscow.

Vendler, Z. (1957). Verbs and Times. *The Philosophical Review* 66:2, pp. 143-160.

Waltke B. K. & M. P. O'Connor (1990). *An Introduction to Biblical Hebrew Syntax*. Eisenbrauns.

Yasin, A. (2014). Cognate Accusatives: lost in translation. *European Scientific Journal* 10:26, pp. 327-353.

Norwegian compounds: the category of non-heads
Ragnhild Eik

In this paper, the inner structure of compounds is discussed, based on an investigation of left-hand elements of compounds in Norwegian. In Dutch and Swedish, left-hand elements of compounds have been analysed as category-less roots without any word-class features (i.e. noun, verb, adjective). I investigate whether this analysis can be extended to Norwegian compounds. A careful examination of left-hand-elements and their interaction with linking elements in Norwegian reveals that at least some categorization is necessary. In Norwegian, non-heads cannot be bare roots. The paper thus contributes to our understanding of the buildings-blocks of compounds in different languages.

1. Introduction

The category of a lexical element can be determined by the element's distribution and paradigmatic behaviour (ten Hacken 1994:86). In a language such as Norwegian, for example, a verb is recognizable by its ability to take arguments and tense-marking. Non-heads of compounds are interesting because they do not have any such context to help us determine their category. Consider the English compound in (1).

(1)   surf board

In the Germanic languages, compounding is right-headed. In the compound in (1), *board* is the head, and *surf* is the non-head. This paper is concerned with non-heads, such as *surf*. Is *surf* in this context a noun or a verb? Or perhaps *surf* does not have a category when used as the non-head of a compound. An analysis where non-heads do not belong to a lexical category is available in a decompositional framework such as Distributed Morphology (e.g. Harley & Noyer 1999; Embick & Noyer 2007) where categories are structurally defined and syntax manipulates category-less roots.

   Regarding the compound in (1), it has been argued that non-heads in English are always categorized (De Belder 2017; see also Harley 2009). Categorization is necessary to capture the claim that English does not allow verbal non-heads in compounds.[1] In other words, *surf* in

---

[1] But see Bauer et al. (2013) who challenge this empirical generalization.

(1) would have to be analysed as a noun. For other Germanic languages, notably Dutch and Swedish, De Belder (2017) and Josefsson (1998) argue respectively that non-heads are not formally categorized as nouns, verbs or adjectives. Instead, non-heads in these languages can be bare forms or roots[2].

In this paper, I address the linguistic nature of non-heads of compounds. More specifically, I ask whether compound-non-heads in Norwegian are formally categorized as belonging to a word-class (i.e. noun, verb, adjective), or whether they are categoryless roots. I investigate this from the point of view of the decompositional models of Distributed Morphology and the exoskeletal model of Borer (2005, 2013).

The question regarding the nature and size of non-heads will be looked at from two different angles. First, I ask whether there is any evidence for categorization in non-heads. If non-heads can be shown to have a category, then it follows from the framework that they are not roots. In order to answer this, I look at the interaction between categories and linking elements in Norwegian compounds. Second, I investigate whether non-heads in Norwegian appear to have internal complexity. Roots are by definition simplex objects. If non-heads are internally complex, then they are not roots.

In theory, the two tests for categorization should converge on the same answer, but we will see that this is not always the case. Most of the data favour an approach where non-heads in Norwegian are categorized, and I will propose such an analysis. On this point, the proposed analysis of Norwegian reaches the opposite conclusion of previous studies on Dutch and Swedish.

Before diving into the investigation of non-heads, some empirical clarification is needed. In all three languages, Dutch, Norwegian and Swedish, there are compounds whose non-heads are clearly nominalized. Some examples from Norwegian are given in (2a-c).

(2)　　a.　forsk-**ing**-s-resultat
　　　　　research-NMLZ-LINK-result
　　　　　'research result'

　　　　b.　sunn-**heit**-s-teikn
　　　　　health-NMLZ-LINK-sign
　　　　　'sign of good health'

　　　　c.　kjær-**leik**-s-brev
　　　　　dear-NMLZ-LINK-letter
　　　　　'love letter'

In these examples, the presence of overt nominalizers indicates that the non-heads in these cases are really nouns. Here, there is no question. What I am concerned with in the current paper is how to deal with compounds where no overt categorizer is present, as in examples (3a-c).

---

[2] De Belder (2017) uses the term 'root', whereas Joefsson (1998) uses 'bare stem'. The difference in terminology is a reflection of the time in which the contributions were written. In both analyses, however, non-heads are bare forms without any syntactic category. I will use the term 'root'.

(3)	a. blå-bær	b. sov-e-pille	c. dag-bok
	blue-berry	sleep-LINK-pill	day-book
	'blueberry'	'sleeping pill'	'diary'

Are the non-heads in (3) categorized by a zero-categorizer, or are they roots?

This paper is organized as follows. In section 2, I provide some theoretical background on categorization and roots. In section 3, I look at linking elements and their distribution in Dutch and Swedish. Section 4 then turns to the relationship between linkers and non-heads in Norwegian. The argument presented in this section is simple: non-heads of different categories take different linkers. Therefore, non-heads must be categorized. In section 5, I ask whether non-heads have the makeup of roots. In section 6, I propose an analysis of non-heads in Norwegian. Section 7 concludes the paper.

2. Some theoretical background on roots and categories

An important assumption in syntactic approaches to word-formation is that words are built in a way similar to phrases. Selkirk's (1982) *The syntax of words* is an early investigation of this approach. In her analysis of compounds, Selkirk (1982:22) proposes the structure in (4).

(4)

N [+ plur]
N   N [+ plur]
apron   strings
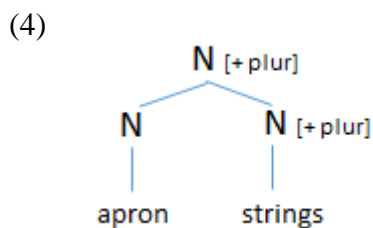
In Selkirk's analysis, and similar analyses at the time, lexical elements already have their category when entering syntax. From this point of view, the question of whether non-heads are categorized is not a pertinent one. It follows from the grammatical architecture that they are.

This is different in more recent frameworks such as Distributed Morphology (DM; Embick & Noyer 2007, Embick & Marantz 2008, Embick 2015) and Borer's exoskeletal syntax (Borer 2005a,b, 2013). In these frameworks, categories are assigned syntactically. In DM, a noun is composed of a root and a little n-head, a verb is composed of a root and a little v-head, and an adjective is composed of a root and a little a-head. The categorizing heads can be realized by derivational suffixes or they can have zero-realizations. In principle, the same root can appear in all three environments, see (5)-(7).

(5)

nP
Vapron   n

(6)

vP
Vrun   v

(7)

aP
Vblue   a
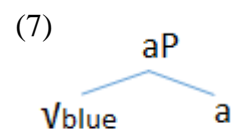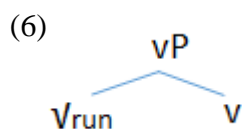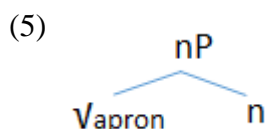
We see now how the non-head of a compound can be uncategorized. A compound with a root as a non-head would have a structure as in (8), and a categorized non-head would look like (9).

(8)

√apron   *string*

(9)

nP   *string*

√apron   n

In addition to providing the root with a category, the categorizing head is hypothesized to play a semantic role. According to Arad (2003), roots are semantically underspecified. A root is assigned a specific, idiosyncratic semantic interpretation in the context of the first categorizing head. The semantic interpretation which is assigned in the context of the first categorizer, is carried along in further derivation (the so-called *Marantz/Arad hypothesis*, Arad 2003, 2005, Marantz 2001, 2007, see also Samioti and Anagnostopoulou 2014 for discussion).

This view of the semantics of roots is not immediately compatible with the hypothesis that non-heads of compounds are bare roots. According to Arad's view, roots are semantically underspecified. A non-head, on the other hand, seems to have more semantic content than the underspecified meaning that roots have in Arad's system. A solution proposed by Josefsson for Swedish is that non-heads are interpreted prototypically. In a compound such as *bokklubb* 'book club', the non-head *bok* is not formally categorized as a noun, but since *bok* is usually used as a noun, we interpret it as a noun (Josefsson 1998:56).

Another way of proceeding has been to challenge the claim that categorizers have this role in negotiating meaning. Borer (2014) argues that Arad's view on idiosyncratic meaning is too strict. Borer provides a number of examples, some of which are given in (10), to show that idiosyncratic meaning is not always established in the context of the first categorizer. Instead, it can be established further out in the word.

(10)   a. react-ion$_{\text{NMLZ}}$-ary$_{\text{ADJ}}$
       b. civil-iz$_{\text{VBLZ}}$-ation$_{\text{NMLZ}}$
       c. except-ion$_{\text{NMLZ}}$-al$_{\text{ADJ}}$

All of the examples in (10) are ambiguous between a compositional and a non-compositional reading. For example, *civilization* can refer to the process of *civilizing* or it can have a culture/society-reading. Only the first reading carries with it the semantic interpretation that was established by the first categorizer. In other words, the meaning of both *civil* and *civilize* is present in the process-reading but not in the culture/society-reading. This shows that idiosyncratic meaning can be assigned to bigger structures than what is assumed in Arad's system. According to Borer, then, it is not the case that meaning is necessarily established by the first categorizer and carried along in further derivation (see Marantz 2013 for discussion of such examples).

Based on Borer's arguments, Scher & Nobrega (2014) conclude that the semantic argument for categorization in non-heads can be abandoned. Non-heads do not have to be categorized in order to give the right semantic interpretation. This move allows them to analyse non-heads of compounds as roots in Brazilian Portuguese. That non-heads are roots has been proposed at least for Dutch (De Belder 2017), Swedish (Josefsson 1998), Greek

(Iordachioaia et al. 2017) and Brazilian Portuguese (Scher & Nobrega 2014). In the following, I will compare Norwegian to Dutch and Swedish. These three languages are closely related and their compounds are very similar. Yet, the conclusions reached are different.

3. Linkers and categories

In compounding, there is very little morphological context to reveal what category the non-head might belong to. The question, then, is how we can determine whether such non-heads are categorized or not. Fortunately, there are cases where there is some overt morphological material in non-heads. Dutch, Swedish and Norwegian all make use of linking elements in compounding, and I will argue that these elements give us some insight into the morphological nature of non-heads.

3.1. Linkers

A linking element is a semantically empty element appearing between the head and non-head of a compound (Lieber & Stekauer 2009:13), e.g. *-e-* in the Norwegian example in (11a). In the Germanic languages, linkers are mainly derived from genitive and plural markers.

Although the linker can be said to appear *between* the non-head and the head, it is clear that it is closer to the non-head than to the head. There are two reasons for making this claim. First, in coordination the linker stays with the non-head, as illustrated with examples from Norwegian in (11b-c).

(11) a. katt-e-mat
         cat-LINK-food
         'cat food'

     b. katt-e   og   hund-e-mat
         cat-LINK and dog-LINK-food
         'cat and dog food' (i.e. cat food and dog food)

     c. katt-e-mat   og *-e-drikke
         cat-LINK-food and -LINK-drink
         intended: 'cat food and cat drink'

Second, the linker is selected by the non-head. The same non-head always appears with the same linker, regardless of the head. In Norwegian, for example, *katt* 'cat' as a non-head always takes an *e*-linker (12a-c), and *fred* 'peace' as a non-head always takes an *s*-linker (13a-c).

(12) a. katt-e-mat   b. katt-e-jakt   c. katt-e-grå
         cat-LINK-food   cat-LINK-hunt   cat-LINK-grey
         'cat food'   'cat hunt'   'cat grey'

(13) a. fred-s-arbeid   b. fred-s-due   c. fred-s-pipe

peace-LINK-work   peace-LINK-dove   peace-LINK-pipe
'peace work'   'peace dove'   'peace pipe'

3.2. Linkers in Dutch and Swedish are sensitive to noun-hood

In Dutch and Swedish, linkers appear on certain nominal non-heads, see (14a-b) and (15a-b) respectively.

(14) Dutch (De Belder 2017:140)   a. stol-s-rygg
     a. varken-s-hok   chair-LINK-back
       pig-LINK-pen   'back of a chair
       'pig's pen'

     b. kat-en-luik
       cat-LINK-panel   b. kyrk-o-musik
       'cat door'   church-LINK-music
                              'church music'

(15) Swedish (Josefsson 1998:61, 65)

In Dutch (16a-b) and Swedish (17a-b), linkers do not appear with non-heads of other categories.

(16) Dutch (De Belder 2017:141, 143)   (17) Swedish (Josefsson 1998:57)
     a. snel-trein   a. röd-färg
       fast-train   red-paint
       'high-speed train'   'red paint'

     b. slaap-pil   b. skriv-maskin
       sleep-pill   write-machine
       'sleeping pill'   'type writer'

The observation that linkers only appear on nominal non-heads suggests that linkers are sensitive to the category of the non-head in these languages. This is in part what leads De Belder (2017) to claim that Dutch compounds with linkers have categorized non-heads, whereas compounds without linkers have bare roots as non-heads.

    Josefsson (1998), although noting that linkers are sensitive to noun-hood in Swedish, does not incorporate this feature into her analysis of Swedish compounds. The non-heads in (15) are not formally categorized as nouns in her analysis[3].

---

[3] Josefsson's reasoning is that this type of linker is lexical in nature and not productive. Based on similar observations for German, Neef (2015) proposes to analyse this type of linker as part of a compound stem. Non-heads are then a case of stem-allomorphy, where the insertion of a compound stem is licenced by compound structure. Although Josefsson does not make her proposal explicit, an analysis along the lines of Neef (2015) seems to fit well with her reasoning.

4. Linkers and categories in Norwegian

In Norwegian, compounds are built similarly to Dutch and Swedish. One important difference, however, is how the languages treat verbal non-heads.

4.1. Nouns and verbs take different linkers in Norwegian

Looking first at nominal non-heads, Norwegian behaves likes Swedish and Dutch. Nominal non-heads receive an *e*-linker (18a), an *s*-linker (18b), no linker (18c) or marginally some other linker (18d). The linking element is selected by the non-head (Iversen 1924, Faarlund et al 1997:61-90), as was also shown in (12) and (13).

(18) a. katt-e-mat   day-book
    cat-LINK-food   'journal'
    'cat food'
   b. fred-s-pipe   d. ferd-a-folk
    peace-LINK-pipe   journey-LINK-people
    'peace pipe'   'travellers'
   c. dag-bok

Most linkers in Norwegian are historically derived from genitive markers. Modern Norwegian no longer has the genitive case.

 In Norwegian, unlike Dutch and Swedish, a linker also appears after verbal non-heads. Compare the examples from Dutch, Swedish and Norwegian below.

(19) Dutch	(20) Swedish	(21) Norwegian
slaap-pil	skriv-maskin	skriv-e-bok
sleep-pill	write.machine	write-LINK-book
'sleeping pill'	'type writer'	'note book'
(De Belder 2017:143)	(Josefsson 1998:57)	(Vinje 1973:63)

Verbal non-heads take linkers in Faroese (*-i, -u* and *-a*, Thráinsson et al. 2004:207) and Icelandic (*-i* and *-u,* Harðarson 2016:8) as well. The Norwegian *-e*-linker on verbal non-heads has developed from an *–i* corresponding to the verbal linker in Faroese and Icelandic (see Conzett 2016).

 Whereas Norwegian nominal non-heads may take an *e*-linker or an *s*-linker (and occasionally some other linker), verbal non-heads only take *e*-linkers. An *s*-linker is impossible with a verbal non-head. This is shown in (22)-(24) below.

(22)	(23)	(24)
a. *skriv-s-bok	a. *sitt-s-plass	a. *tenk-s-pause
b. skriv-e-bok	b. sitt-e-plass	b. tenk-e-pause
write-LINK-book	sit-LINK-place	think-LINK-pause
'note book'	'seat'	'pause for thought'

It is difficult to see how this generalization can be made without referring to the category of the non-head. In order for a speaker of Norwegian to successfully create a new compound, she must know the category of the non-head. I interpret this to mean that non-heads have a category.

The data presented in this section provides the first indication that verbal and nominal non-heads are treated differently in Norwegian, and hence that they are categorized. The distribution of linkers can be summarized as in (25).

(25)
      Verbal non-head: *-e*
      Nominal non-head: *-e, -s, (-a, -ar, -er, -en)*

Adjectival non-heads do not take linkers.

Before we can firmly conclude that non-heads in Norwegian compounds are formally categorized, there are some additional factors to consider. I turn to those in the next subsection.

4.2. Is the -e on verbs really a linker?

One possible objection to the interpretation of the data presented in the previous section is that the linker on verbal non-heads could in fact be the infinitival ending, which is also realized as *-e* in some dialects, such as the Oslo dialect. This interpretation would challenge the claim that verbal and nominal non-heads take different linkers.

Such an account is not available for a large number of dialects where the infinitive has other realizations. Compare the infinitive in the a.-examples (*read*-INF) of (26) to (28) with the compound-form in the b.-examples (*read*-LINK-*book* 'textbook') for the following Norwegian dialects.

(26) Oslo dialect	(27) Hardanger dialect	(28) Salten dialect
a. les-e	a. les-a	a. læs
b. les-e-bok	b. les-e-bok	b. læs-e-bok

Only the Oslo dialect supports an analysis in which verbal non-heads make use of the infinitive.

4.3. Verbal e-linkers and nominal e-linkers

I have claimed that linkers are sensitive to the category of the non-head, and interpret this to mean that non-heads are categorized. There is yet another way to interpret the data presented thus far. As was shown in (18) and (25), some nouns take *e*-linkers, just as verbs do. Based on such examples, an alternative hypothesis is that linkers are sensitive not to category, but to some other classification. For example, there could be a set of roots which select *e*-linkers and

another set of roots which select *s*-linkers. Verbal roots could all 'happen to be' in the *e*-selecting group. In this section, I argue that that hypothesis is not correct.

For this discussion, it is relevant to distinguish between two classes of nouns: strong nouns and weak nouns. In Germanic, the distinction between so-called strong and weak nouns is, or has been, relevant for nominal declension classes (see Berg forthcoming on this distinction in West Nordic). In Modern Norwegian, strong nouns end in a consonant in their bare form and weak nouns end in an unstressed vowel in their bare form.

The observation I will make is that even if we postulate *e*-selecting roots, we have to distinguish between three groups of such roots: verbal *e*-selecting roots, strong nominal *e*-selecting roots and weak nominal *e*-selecting roots. The set of non-heads that takes *e*-linkers do not behave uniformly, so the relevant generalizations must still make reference to the category of the non-head. For this reason, the alternative hypothesis sketched above does not take us any further.

In order to see that the alternative hypothesis will not fare any better, I first compare strong nouns on the one hand to verbs and weak nouns on the other. Next, I compare verbs and weak nouns, and we will see that a three-way distinction between these classes is needed.

Strong nouns take a number of different linkers (*-e, -s, -a, -ar, -er, -en, etc.*). The *e*-linker on strong nouns is lexically selected, it is unpredictable and its productivity is very limited. Furthermore, the *-e* on strong nouns is not carried over in complex compounding. Compare the simple compound in (29) and the complex compound in (30), with the Norwegian strong nominal non-head *mjølk* 'milk'.

(29) *Simple compound: strong noun*   (30) *Complex compound: strong noun*
     [mjølk-e]-glas   [sur-mjølk-s]-glas
     milk-LINK-glass   sour-milk-LINK-glass
     'milk glass'   'glass for cultured milk'

Mjølk is a strong noun, which we can see because it is consonant-final. When *mjølk* is the first element of a compound it takes an *e*-linker, as in (29). When *mjølk* is the head of a complex non-head, it takes an *s*-linker, as in (30).

Strong nouns differ from weak nouns and verbs in a number of ways. Weak nouns and verbs only occur with one linker: *-e.* This *e*-linker is predictable and productive, so its distribution is rule-based.[4] Furthermore, the linker on weak nouns and verbs is carried over in complex compounding. The pattern is shown in (31)-(34).

(31) *Simple compound: weak noun*   (32) *Complex compound: weak noun*
     skjort-e-strykar   [fin-skjort-e]-strykar
     shirt-LINK[5]-ironer   nice-shirt-LINK-ironer
     'person who irons shirts'   'person who irons nice shirts'

---

[4] The productivity of linkers on verbs can be seen, for example, in borrowings from English. English verbs are frequently borrowed into Norwegian, and they always receive an *e*-linker in compounding.
    i.   **chill**-e-søndag   ii.   **sav**-e-vane
        chill-LINK-Sunday   save-LINK-habit
        'relaxing Sunday'   'habit of saving f.ex. in a computer game'
[5] For some dialects, the status of this - as a linker is unclear because the compound form is homophonous with the free form. For example, in the Oslo dialect, we find *en skjorte* 'a shirt' and *skjortestryker* 'shirt ironer'.

(33) *Simple compound: verb*   (34) *Complex compound: verb*
      les-e-strategi   [hurtig-les-e]-strategi
      read-LINK-strategy   fast-read-LINK-strategy
      'reading strategy'   'speed-reading strategy'

In (31)-(34), the non-head receives the same linker, *-e*, in both simple and complex compounding. At first glance, then, it can seem as though we have an alternative classification where verbs and weak nouns pattern together on the one hand, against strong nouns on the other. If so, the category of the non-head is not the determining factor for the choice of linker. Linkers could instead be sensitive to some other property, such as the phonology of the non-head.

What is needed in order to disprove this hypothesis is a variety of Norwegian where verbal non-heads and weak nominal non-heads are also treated differently in compounding. This is exactly what speakers of the Gudbrandsdalen dialect report. In this dialect, there is a clear three-way distinction between strong nouns, weak nouns and verbs as non-heads, as illustrated by examples (35)-(40).

(35) *Simple compound: strong noun*   (36) *Complex compound: strong noun*
      [mjølk-e]-glas   [sur-mjølk-s]-glas
      milk-LINK-glass   sour-milk-LINK-glass
      'milk glass'   'glass for cultured milk'

(37) *Simple compound: weak noun*   (38) *Complex compound: weak noun*
      skjort-strykar   [fin-skjort]-strykar
      shirt-ironer   nice-shirt-ironer
      'person who irons shirts'   'person who irons nice shirts'

(39) *Simple compound: verb*   (40) *Complex compound: verb*
      les-e-strategi   [hurtig-les-e]-strategi
      read-LINK-strategy   fast-read-LINK-strategy
      'reading strategy'   'speed-reading strategy'
                   (Norwegian Gudbrandsdalen dialect, p.c. P.M. Kveen, P.E. Myren-Svelstad)

In this dialect, strong nouns and verbs behave as in other dialects (compare (35)-(40) to the other Norwegian examples in (29)-(34)), but weak nouns never take a linker, as shown in examples (37) and (38).

At least in this dialect, it is not the case that verbal roots and weak nominal roots form a class, such that linkers could be sensitive to some other classification that does not make reference to lexical category. Speakers must know the category (and declension) of the non-head in order to form new compounds.

As far as I can tell, there is no alternative classification of roots which can predict the behaviour of non-heads and linkers. The data strongly suggests that non-heads in Norwegian are categorized.

4.4. Summarizing the evidence from linkers

The previous sections have shown that verbal and nominal non-heads are not treated the same way in compounding. This is true, even for verbs and weak nouns, which in some cases appear to be treated similarly. In conclusion, following the reasoning whereby linkers are sensitive to lexical categories, verbal non-heads are categorized as verbs, and nominal non-heads are categorized as nouns.

5. Internal complexity in apparently simple non-heads

The previous section investigated whether there is evidence for categorization in non-heads. In this section, I look at non-heads from a different point of view: Do non-heads of compounds fit the description of roots? We will see that the data points in different directions, and the answer to the question appears to be 'In size, yes. In internal complexity, no.'

5.1. Overt verbal and adjectival categorizers are disallowed

There are some restrictions on the morphological size of elements used as non-heads of compounds. De Belder (2017) points this out for Dutch, and notably, the same restrictions hold in Norwegian.

De Belder's (2017) observation is that overt adjectival and verbal derivational morphology is prohibited in non-heads. She gives the following examples (41-42) from Dutch (De Belder 2017:144).

(41) No overt little a-head
    a. *spar-zaam-attitude
       save-ADJ-attitude

    b. *eet-baar-datum
       eat-ADJ-date

    c. *vet-ig-dieet
       fat-ADJ-diet

(42) No overt little v-head
    a. *?be-plant-seizoen
       VBLZ -plant-season

    b. *?menstru-eer-pijn
       menstru-VBLZ-pain

    c. *?ont-vlam-tijd
       VBLZ-flame-time

The examples in (41) and (42) contrast in acceptability with compounds where non-heads are verb-like and adjective-like, but there is no overt categorizer (De Belder 2017:141, 145).

(43) snel-trein
    fast-train
    'high-speed train'

(44) eet-datum
    eat-date
    'eating date'

For De Belder, the best way to account for the impossibility of overt verbal and adjectival categorizers is to say that non-heads in these cases are roots. If the non-heads in (43) and (44) were categorized by little categorizing a- and v-heads, then we would expect overt categorizers to be possible as well. "After all", she points out "null categorial heads and overt categorial heads are only distinguished at PF given late insertion. Structurally, they are

identical." (De Belder 2017:144). In other words, for De Belder, *snel* and *eet* are not formally speaking an adjective and a verb in these examples. They are bare roots.

Turning now to Norwegian, we find the same restrictions as in Dutch. Overt verbalizing and adjectivizing morphology is clearly dispreferred. Overt nominalizing morphology, however, is perfectly fine.

Example (45) shows that a wide range of overt nominalizers are allowed in non-heads. Examples (46) and (47) show that overt adjectival and verbal suffixes are not nearly as acceptable[6]. Note that there is very little verbalizing morphology in Norwegian, making the restriction on verbalizers less clear. Note also that elements marked with a star can perhaps be coerced into acceptability by turning them into phrasal compounds. The important observation, however, is that they are clearly worse than the corresponding nominalized non-heads and root-sized non-heads.

(45)   Overt nominalizers in Norwegian non-heads: no restrictions

---

[6] In assessing these examples, I have asked for judgements from six native speakers of Norwegian. This provides a preliminary impression of their acceptability, but further investigation may be needed to corroborate this impression. In addition, searches in the corpus 'Ordbanken' give only very few hits for adjectival and verbal derivational morphology in non-heads. These will be commented on in footnotes.

a. *-sjon* konstruk-sjon-s-arbeid
    construct- NMLZ- LINK-work
    'construction work'

b. *-ing* forsk-ing-s-result
    research-NMLZ-LINK-result
    'research result'

c. *-dom* mann-dom-s-prøve
    man-NMLZ-LINK-test
    'manhood test'

d. *-skap* far-skap-s-test
    father- NMLZ-LINK-test
    'paternity test'

e. *-het* sunn-het-s-tegn
    health-NMLZ-LINK-sign
    'healthy sign'

f. *-leik* kjær-leik-s-brev
    dear-NMLZ-LINK-letter
    'love letter'

g. *-itet* lojal-itet-s-teikn
    loyal-NMLZ-LINK-sign
    'loyalty sign'

h. *-ar* frels-ar-son
    save-NMLZ-son
    'saviour son'

i. *-nad* lov-nad-s-ord
    promise-NMLZ-LINK-word
    'word of promise'

j. *-else* dann-else-s-reise
    create-NMLZ-LINK-journey
    'educational journey'

k. *-sel* hør-sel-s-plage
    hear-NMLZ-LINK-problem
    'hearing problem'

(46) Over verbalizers in Norwegian non-heads[7]
    a. *-er* *konstru-er-(e)-arbeid[8]
        construct-VBLZ-(LINK)-work

    b. *-iser* *nominal-iser-(e)-prosess
        nominal-VBLZ-(LINK)-process

    c. *-n* *gul-n-(e)-grad
        yellow-VBLZ-(LINK)-degree

---

[7] It is also impossible to use a particle verb as a non-head in Norwegian, cf. *an.komme-tid,*'arrive time' i.e. 'arrival time', *ned.kjøle-apparat 'cool device'*. In order for particle verbs to become acceptable as non-heads, they must be nominalized, cf. *ankomst-tid* 'arrival time', *nedkjølings-apparat* 'cooling device'. In contrast, De Belder (2017) reports that particle verbs are fine as non-heads in Dutch.

[8] Counter-examples: *barbermaskin* 'electric razor' and *justervesenet* 'the Norwegian metrology service'. In these cases, the verbalizer could be analysed as part of the root, since there is no independent *barb-* or *just-*.

(47) Overt adjectivizers in Norwegian non-heads[9]
    a. *-leg*   *?barn.s.leg-haldning
               child- LINK-ADJ-attitude
               intended: 'childish attitude'

    b. *-ig*   *?sjølv-stend-ig-følelse[10]
               self-stand-ADJ-feeling
               intended: 'feeling of independence'

    c. *-ete*   *?bråk-ete-barn
               noise-ADJ-child
               intended: 'noisy child'

    e. *-bar*   *?hald.bar.dato
               hold-ADJ-date
               intended: 'expiration date'

    f. *-sam*   *?spar-sam-haldning
               save-ADJ-attitude
               intended: 'attitude for saving'

All of the non-heads in (46)-(47) must be nominalized in order to become acceptable.

From the data presented in (45) to (47), the generalization seems to be that Norwegian allows nominalized and 'root-sized' elements as non-heads. On the one hand, then, we have an indication that verbal and adjectival non-heads must be 'root sized'. On the other hand, we saw in section 4 that there is reason to believe that non-heads are categorized. The next pieces of data also favour an approach where non-heads are categorized.

5.2. Nominal vs. verbal roots

A key idea behind the root-hypothesis of Distributed Morphology is that one and the same root may be present in words of different categories. For example, the noun *dream* and the verb *dream* both contain the same root √DREAM with nominalizing and verbalizing structure.

In some cases, the verbal and nominal versions of a root have slightly different phonological realizations. This is the case with Norwegian (Nynorsk) *draum/drøym* 'dream'. Both the nominal version *draum* and the verbal version *drøym* can be used as the non-head of a compound, as shown in (48). Further examples of this phenomenon are given in (49) and (50).

---

[9] Most overt adjectivizers are perfect when the head of the compound is *gjere* 'do, make', e.g. *ufarleggjere* 'render harmless'. This appears to be a kind of incorporation structure, and it is not clear to me that it should be analysed as ordinary compounding.

[10] A possible non-head with *-ig* is *heldiggris* 'lucky pig'. In this case, the suffix is best analyzed as part of the root.

(48)

    a. draum-e-hus
      dream[N]-LINK-house
      'dream house' (e.g. a house made of dreams)

    b. drøym-e-hus
      dream[V]-LINK-house
      'dreaming house' (e.g. a house in which you dream)

(49)

    a. song-stund
      song[N]-time
      'sing-song'

    b. syngj-e-stund
      sing[V]-LINK-time
      'time for singing'

(50)

    a. tank-e-tank
      thought[N]-LINK-tank
      'tank filled with thoughts'

    b. tenk-e-tank
      think[V]-LINK-tank
      'a tank for thinking'

The verbal and nominal version of the roots contribute different semantics to the compound. If we follow what I take to be the standard assumption, that both the noun *draum* and the verb *drøym* contain the same semantically and phonologically underspecified root with additional categorizing structure, then the examples above indicate that non-heads are categorized. If so, they are morphologically complex and they are not roots.

5.3. Causative vs. anti-causative roots

Some verbs have a causative and an anti-causative version. In Norwegian, both versions can participate in compounding. The causative and anti-causative version make different semantic contributions to the compound, as demonstrated in (51) to (53).

(51)

    a. sitt-e-streik
      sit[ANTIC]-LINK-strike
      'sit in'

    b. sett-e-potet
      set[CAUS]-LINK-potato
      'seed potato'

(52)

    a. ligg-e-stol
      lie[ANTIC]-LINK-chair
      'reclining chair'

    b. legg-e-tid
      lay[CAUS]-LINK-time
      'bedtime'

(53)

    a. synk-e-ferdig
      sink[ANTIC]-LINK-finished
      'ready to sink' (e.g. a boat in bad shape)

    b. senk.e.kjøl
      sink[CAUS]-LINK-keel
      'drop keel'

Verbs such as these are often analysed as stem alternations of a common root (see e.g. Alexiadou, Gehrke and Schäfer 2014). The amount of functional material in the structure determines which phonological realization is used. If we follow this approach, the non-heads in (51) to (53) are morphologically complex and contain at least a little v-head.

6. Towards an analysis of non-heads

The general finding of the previous sections is that non-heads must be categorized. Most attention has been given to verbal non-heads since these distinguish Norwegian from Dutch and Swedish. I propose the structure in (54) for verbal non-heads.

(54)

```
              v        LINK
                        -e
        Vskriv   v
                 ø
```
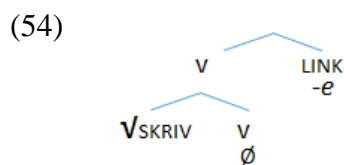
I assume corresponding structures for nominal and adjectival non-heads.

As we have seen throughout the paper, there are several reasons for assuming that non-heads in Norwegian are categorized as belonging to a word-class. First, the choice of linker in Norwegian appears to be sensitive to the category of the non-head. Second, in cases where there are several versions of the same root – a nominal and a verbal version as in (45) to (47) or a causative and an anti-causative version as in (48) to (50) – both versions can appear as the non-head of a compound. If we assume an analysis where such pairs are realizations of a common root with different layers of functional structure, then it follows that non-heads have some structure. In other words, they are not roots.

On the other hand, the observation that verbal and adjectival non-heads cannot be morphologically complex would favour a root-analysis for non-heads. This is what De Belder (2017) proposes based on corresponding restrictions in Dutch. I do not have a final explanation of this restriction, but I sketch here three factors that are likely to be relevant and that should be considered in future work.

First, the contrast between overt and internal complexity in adjectival and verbal non-heads is reminiscent of another restriction on non-heads. It is well known that English allows irregular plural nouns but not regular plural nouns as non-heads of compounds (see e.g. Siddiqi 2009). This is shown in (55). The pattern is less common in Norwegian, but an example is given in (56a).

(55)   (56)
    a. lice-infested   a. born-e-born
    b. *rat-s-infested   child[PL]-LINK- child[PL]
                                     'grand children'

                                     b. *bøk-er-hylle
                                       book-PL-shelf

It could be that the restriction on overt morphological material in non-heads is more general. The same restriction could be ruling out overt adjectivizers, verbalalizers and inflectional morphemes such as plural marking. If so, what is in need of explanation, is perhaps rather why overt nominalizers are allowed when other morphological material is not.

Second, a similar restriction on derivational suffixes is found in conversion. In English, a simple root can easily be converted between noun and verb. For example, we have both the verb *to form* and the noun *a form.* Conversion, or zero-derivation, cannot apply when there is overt derivational material. We cannot zero-derive a noun **a formalize* from the verb *to formalize.* Nor can we zero-derive a verb **to formalization* from the noun *a formalization.* Different explanations have been proposed to account for this (see e.g. Borer 2003). Whatever explanation we assume for conversion can probably be extended to restrictions on overt material in compounding. Just as derivational suffixes and processes have restrictions on what elements they can take as input, so does the compounding process have restrictions on what it can take as input. There could be a common explanation for the restriction on overt morphological material in conversion and compounding. This brings us to the third consideration.

Third, there is a thought-provoking proposal by Borer (2013) in which there is a compound frame which makes the non-head N$_{\text{oun}}$-equivalent. In Borer's system, this means that non-heads can be roots or already nominalized elements. Another way of putting this is that the compound frame converts the non-head into a noun. Could non-heads in Norwegian be N-equivalent[11]? The immediate answer should be no. We have seen a range of examples showing that verbal non-heads are categorized as verbs. On the other hand, as has already been pointed out, the Norwegian verbal linker has developed from an *i*-linker such as the one we find in Faroese and Icelandic. The authors who have considered this linker all hint to its possible status as a nominalizer (see Harðarson 2016, Conzett 2016). I am hesitant to say that the verbal linker *is* a nominalizer. There is no *–e*-nominalizer outside of compounds that can be used to create nouns from verbs. On the other hand, to my knowledge, all linking elements in Germanic have developed from nominal inflection. This is the case with the most common linkers developed from genitive and plural markers. It is also the case with the *-ar*-linker in Norwegian, developed from the *nomen agentis* marker, corresponding to English *-er,* and it is the case with the diminutive suffix in West-Frisian, which is also used as a linking element (Hoekstra 1996, cited in Fuhrhop & Kürchner 2015). It could very well be the case, then, that linking elements are nominal in nature. With verbal non-heads, the linker could make the non-head 'nominal enough'. If so we have a possible explanation for the restrictions on complexity in non-heads: Non-heads must be either nominal or roots (which can be made noun-equivalent by the structure). In Norwegian, verbal non-heads are made nominal by the linker. In Dutch and Swedish, verbal non-heads are roots. This does not explain why it should be the case that verbal non-heads in Norwegian need a linker, while they can be roots in

---

[11] De Belder (2017) considers the option that non-heads are made N-equivalent and rejects it. One of her arguments is that not all non-heads are interpreted nominally. I do not follow this argument. The nominalization process makes the output nominal, not necessarily the input. For example, in a nominalization such as *singing,* the output of the process is a noun. That does not make the input 'sing' nominal. The same can be the case if the compound frame makes the non-head N-equivalent.

Dutch and Swedish. But it does predict two possible strategies for languages to choose from: the Norwegian-Faroese-Icelandic strategy and the Dutch-Swedish strategy.

In this section, I have sketched some possible angles and considerations for future work on restrictions in non-heads. Whichever angle one chooses, there are further details to be worked out, and I leave those for future work.

7. Conclusion

The starting point of this paper was a claim in the literature that non-heads of compounds are bare roots. In the Germanic languages, such an analysis has been proposed for Dutch and Swedish. I have shown that this analysis cannot be adopted for non-heads in Norwegian. My reasoning relied on two different types of data. First, I considered the interaction between non-heads and linkers. Linkers in Norwegian are sensitive to the category of the non-head, and it follows, I argued, that non-heads are categorized. Second, I considered non-heads which are overtly simple, but can be shown to have internal structure. Both types of data indicate that non-heads in Norwegian are not roots. On the other hand, the prohibition of overt adjectival and verbal derivational morphology in non-heads does suggest a restriction on the overt size of non-heads. At the end of the paper, I sketched some possible ways of explaining this. This proposes an interesting area of research in future work. More generally, the paper contributes to our understanding of cross-linguistic variation in the basic building blocks of compounds.

Abbreviations

NMLZ	Nominalizer	V	Verb
LINK	Linker	PL	Plural
VBLZ	Verbalizer	ANTIC	Anticausative
ADJ	Adjectivizer	CAUS	Causative
N	Noun		

Ragnhild Eik
NTNU, Norwegian University of Science and Technology
ragnhild.eik@ntnu.no

References

Anagnostopoulou, E., & Y. Samioti (2014). Domains within words and their meanings: A case study. Alexiadou, A., H. Borer & F. Schäfer (eds.), *The syntax of roots and the roots of syntax*, Oxford: Oxford University Press, pp.81-111.

Arad, M. (2003). Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory* 21, pp. 727-778.

Arad, M. (2005). *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer

Alexiadou, A., B. Gehrke & F. Schäfer (2014). The argument structure of adjectival participles revisited. *Lingua* 149, pp. 118-138.

Bauer, L., R. Lieber & I. Plag (2013). *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press

Berg, I. (forth.) Gender and declension mismatches in West Nordic. Fabrizio, C. & M. Cennamo (eds.), *Historical Linguistics 2013: Selected papers from the 22nd International Conference on Historical Linguistics*, Naples, 27–31 July 2015. Amsterdam: John Benjamins.

Borer, H. (2003). Exo-skeletal vs. endo-skeletal explanations: syntactic projections and the lexicon. Moore, J. & M. Polinsky (eds.) *The nature of explanation in linguistic theory,* Stanford: CSLI Publications, pp. 31–67.

Borer, H. (2005a). *In Name Only*. Structuring sense, vol. 1. Oxford: Oxford University Press.

Borer, H. (2005b). *The normal course of events*. Structuring sense, vol. 2. Oxford: Oxford University Press.

Borer, H. (2013). *Taking Form*. Structuring sense, vol. 3. Oxford: Oxford University Press.

Borer, H. (2014). Derived nominals and the domain of content. *Lingua* 141, pp. 71-96.

Conzett, P. (2016). Ordlaging. Sandøy, H. (ed.), Oslo: Novus, pp. 61-107.

De Belder, M. (2017). The root and nothing but the root: primary compounds in Dutch. *Syntax* 20, pp. 138–169.

Embick, D. & R. Noyer (2007) Distributed morphology and the syntax morphology interface. Ramchand, G. & C, Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press, pp. 289–324.

Embick, D. & A. Marantz (2008). Architecture and Blocking. *Linguistic Inquiry* 39, pp. 1-53.

Embick, D. (2015). *The Morpheme : A Theoretical Introduction* (Vol.31) Berlin/Boston: De Gruyter.

Faarlund, J., S. Lie & K. Vannebo. (1997). *Norsk referansegrammatikk.* Oslo: Universitetsforlaget

Fuhrhop, N. & S. Kürschner (2015). Linking elements in Germanic, Müller, P. O, I. Ohnheiser, S. Olsen, F. Rainer (eds.) *Word-formation : an international handbook of the languages of Europ.,* de Gruyter Mouton, pp. 568-582.

ten Hacken, P. (1994). *Defining morphology: A principled approach to determining the boundaries of compounding, derivation, and inflection.* [PhD thesis]. Universität Basel.

Harðarson, G. (2016). Peeling away the layers of the onion: on layers, inflection and domains in Icelandic compounds. *The Journal of Comparative Germanic Linguistics,* 19:1, pp. 1-47.

Harley, H., & R. Noyer (1999). State-of-the-article: Distributed morphology. *Glot international*, 4:4, pp. 3-8.

Harley, H. (2009). Compounding in distributed morphology. Lieber, R. & P. Stekauer (eds.) *The Oxford handbook of compounding.* Oxford: Oxford University Press.

Hoekstra, E. (1996). Iets over de eerste leden van samenstellingen. *Leuvense Bijdragen* 84, pp. 491–504

Iordachioaia, G., A. Alexiadou & A. Pairamidis (2017) Morpho-syntactic sources for synthetic compounds in English and Greek. *Journal of Word Formation* 1, pp. 47-72.

Iversen, R. (1924). Om sammensetning av substantiv med substantive. *Maal og Minne,* pp. 9-28

Josefsson, G. (1998). *Minimal words in a minimal syntax: word formation in Swedish*. Amsterdam: Benjamins.

Lieber, R. & P. Stekauer (2009). *The Oxford handbook of compounding.* Oxford: Oxford University Press.

Marantz, A. (2001). Words and Things. Ms., MIT & NYU.

Marantz, A. (2007). Phases and words. S. H. Choe et al, (eds.) *Phases in the theory of grammar*. Seoul: Dong In Publisher

Marantz, A. (2013). Locality domains for contextual allomorphy across the interfaces. Matushansky, O., & Marantz, A. P. (eds.), *Distributed Morphology Today: Morphemes for Morris Halle*, 95-115.

Neef, M. (2015). The status of so-called linking elements in German: arguments in favor of a non-functional analysis. *Word Structure* 8.1, pp. 29-52.

Scher, A. P., & V. A. Nóbrega (2015). Unifying neoclassical and stem-based compounds: a non-lexical approach. *Revista Linguística*, 10:1, pp. 74-98.

Selkirk, E. O. (1982). *The syntax of words*. Cambridge, Mass.: MIT Press.

Siddiqi, D. (2009). *Syntax within the word: Economy, allomorphy, and argument selection in Distributed Morphology*, vol. 138, John Benjamins Publishing.

Thráinsson, H., H. P. Petersen, J. í Lon Jacobsen & Z. S. Hansen (2004). *Faroese: An Overview and Reference Grammar.* Tórshavn: Faroese Academy of Sciences

Vinje, F. E. (1973). *Ord om ord: Innledning til en norsk ordlagningslære.* Trondheim: Tapir.

The *de re* reading and the universal quantificational phrase in Mandarin Chinese

Rong Yin

In this paper, I examine the *de re* reading with respect to universal quantificational phrases (i.e., *mei ge xuesheng* 'every student') in Mandarin. I show that data from Mandarin do not support the traditional scope theory of intensionality (see Russell 1905; Montague 1973; Ladusaw 1977; Ogihara 1992, 1996; Stowell 1993) and Keshet's (2011) split intensionality theory for the *de re* reading, but can be explained by a system incorporating world variables (see Cresswell 1990; Percus 2000; Fintel & Heim 2011).

1. Introduction

It is a classic observation that DPs in an embedded clause following a propositional attitude verb can be understood in at least two ways[1]. For instance, in (1) the matrix verb is a propositional attitude verb, namely *to believe*, and the interpretation of the DP *a teacher* in the embedded *that*-clause, which follows the propositional attitude verb *to believe*, is ambiguous.

(1)   John believes that Mary is dating a teacher.

The ambiguity lies in whether the DP *a teacher* is interpreted with regard to the actual world (i.e., the reality) or the attitude holder's belief worlds (i.e., what the attitude holder believes in his/her mind; in this case, what *John* believes in his mind).

In (1), if the DP *a teacher* is interpreted relative to John's belief worlds, (1) can be understood as: John believes that Mary is dating someone, and he believes that the person Mary is dating is a teacher. One can accept this interpretation no matter whether the person Mary is dating is indeed a teacher in reality. Take the following situation as an example. In reality, Mary is actually dating a singer. One day, John happened to see them meeting in a restaurant from another table without having a chance to talk to Mary. By mistake, John believes that person is a teacher. Under this situation, one can say that John believes that Mary is dating a teacher (although in reality, Mary is dating a singer). This interpretation where the DP (in this case, *a teacher*) is interpreted relative to the attitude holder's belief worlds (in this case, *John*'s belief worlds), is called *de dicto* reading. The formal statement of this interpretation is: $\forall w' \in$ Beliefs(John, w_0): $\exists x.$ x is a teacher in w' and Marry is dating x in w'.

---

[1] In this paper, I do not discuss the third reading pointed out by Fodor (1970), where the specificity of the DP is another concern for interpreting the sentence, considering that the specificity is not clear when the DP is a universal quantificational phrase.

In (1), if the DP *a teacher* is interpreted relative to the actual world, (1) can be understood as: John believes Mary is dating a particular person, and this person happens to be a teacher in reality. One can accept this interpretation no matter whether in John's mind, the person Mary is dating is a teacher or not. Take another situation as an example. In reality, Mary is actually dating a teacher. One day, John happened to see the person Mary is dating in the street. By mistake, John believes that person is a dancer. Under this interpretation, one can say that John believes that Mary is dating a teacher (however, in John's mind, the person Mary is dating is a dancer). This is interpretation where the DP *a teacher* is interpreted relative to the reality, is called *de re* reading. The formal statement of this interpretation is: $\exists x.$ x is a teacher in w_0 and $\forall w' \in \text{Beliefs}(\text{John}, w_0)$: Mary is dating x in w'.

In short, the ambiguity is whether the person Mary is dating is a teacher in reality (i.e., *de re* reading), or the person Mary is dating is believed to be a teacher in John's mind (i.e., *de dicto* reading).

This ambiguity is a property of the propositional attitude verbs, and is also found when the DP is not an indefinite phrase. Sentence (2) is an example where the DP is a universal quantificational phrase, the type of DP that I focus on in this paper.

(2)   John believes that Mary saw every horror movie.

The *de re* interpretation for (2) is that for every x, x is a horror movie in the real world, John believes that Mary saw x (although in John's mind, x might not be a horror movie). The *de dicto* reading for (2) is that John believes that for every x, x is a horror movie, Mary saw x (although in reality, x might not be a horror movie, but x is a horror movie in John's mind).

In Section 2, I make a brief introduction to current approaches to explain the *de re* reading: the traditional scope theory of intensionality (see Russell 1905; Montague 1973; Ladusaw 1977; Ogihara 1992, 1996; Stowell 1993), Keshet's (2011) split intensionality theory and a system incorporating world variables (see Cresswell 1990; Percus 2000; Fintel & Heim 2011). In Section 3, I discuss the syntactic properties of the universal quantificational phrase in Mandarin Chinese and how they are interpreted under the propositional attitude verbs. In Section 4, I show that the traditional scope theory of intensionality and the split intensionality theory for the *de re* reading are not adequate to explain the Mandarin data, but the data can be explained by a system incorporating world variables. In Section 5, I discuss problems and conclude.

2.   Current analyses of the de re reading
2.1. The traditional scope theory of intensionality[2]

Under the traditional scope theory of intensionality (STI) for DPs (see Russell 1905; Montague 1973; Ladusaw 1977; Ogihara 1992, 1996; Stowell 1993), when the DP c-commands the intensional verb[3] (i.e., the propositional attitude verb) the sentence receives a *de re* reading; when the intensional verb c-commands the DP, the sentence receives a *de dicto* reading. For sentence (1), the DP in questions is *a teacher*, and the intensional verb is *to believe*.
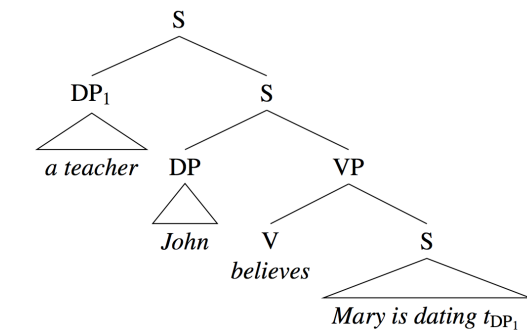
The c-command relationship between the intensional verb *to believe* and the DP *a teacher* is accomplished by applying the Quantifier Raising (QR) operation: the *de re* reading of

---

[2] Section 2.1 and Section 2.3 are based on material from Cable's (2016) unpublished lecture notes.
[3] Intensional verbs are also called intensional operators.

sentence (1) can be derived if one assumes that QR can move the DP *a teacher* across the finite embedded clause at LF, which is shown in (3); the *de dicto* reading of sentence (1) can be derived by moving the DP to a position inside the embedded clause through QR at LF, which is shown in (4). In (3), the DP *a teacher* c-commands the intensional verb *believes*, and thus *de re* reading is received. In (4), the intensional verb *believes* c-commands the DP *a teacher*, and thus *de dicto* reading is derived.

(3)   The LF for the *de re* reading of sentence (1):

(4)   The LF for the *de dicto* reading of sentence (1):
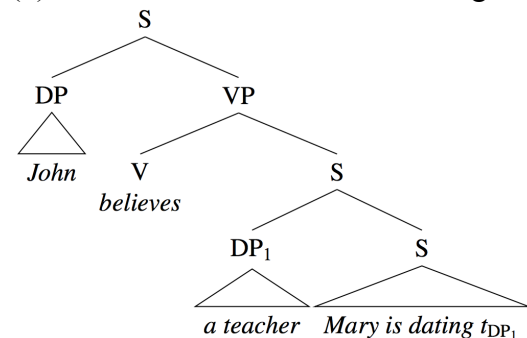
In short, STI provides an analysis for the *de re*/*de dicto* ambiguity by interpreting the c-command relationship between the DP relative to the intensional verb; QR makes the movement of the DP to different positions possible and enables the DP to c-command or to be c-commanded by the intensional verb, with the assumption that QR can move the DP beyond a finite clause boundary.
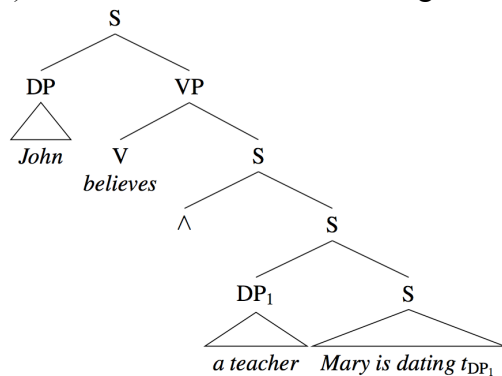
2.2. Keshet's (2011) split intensionality theory

Keshet's (2011) split intensionality theory is different from STI in that the intensionality is divided between an intensional verb ω, like *believes* in sentence (1), and the "up" operator $^\wedge$, which can be inserted freely. The $^\wedge$ operator marks the region of a *de dicto* reading; anything it c-commands receives a *de dicto* reading. In this sense, the *de re* reading of sentence (1) can be derived by assuming (5) as its LF, in which by applying the QR operation, the DP does not need to move to a position outside the finite embedded clause, but a position which is higher than the operator $^\wedge$, which can be inserted within the embedded clause.

(5) The LF for the *de re* reading of sentence (1):

```
                    S
            _____|_____
           DP               VP
          /\          _____|_____
        John         V               S
             believes          _____|_____
                              DP₁           S
                             /\        _____|_____
                        a teacher  ∧          S
                                         _____|_____
                                        Mary is dating t_DP₁
```
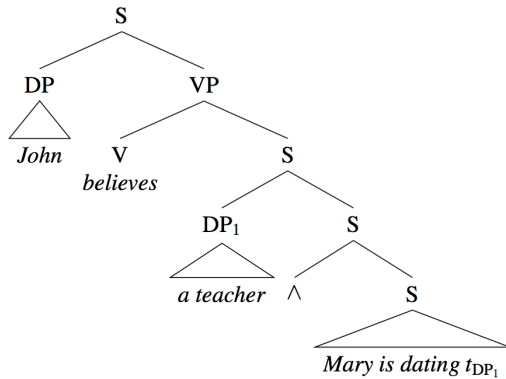
In (5), although the DP *a teacher* is c-commanded by the intensional verb, it c-commands the operator ^, and thus gets the *de re* reading. The LF for the *de dicto* reading is shown in (6), where the DP *a teacher* moves to a position which is c-commanded by the operator ^.

(6) The LF for the *de dicto* reading of sentence (1):

```
                    S
            _____|_____
           DP               VP
          /\          _____|_____
        John         V               S
             believes        _____|_____
                            ∧               S
                                      _____|_____
                                     DP₁           S
                                    /\        _____|_____
                               a teacher   Mary is dating t_DP₁
```

In short, under Keshet's (2011) split intensionality theory, the basic idea of STI still remains in the sense that the *de re*/*de dicto* ambiguity is derived by the c-command relationship between the DP and the intensional operator. However, there is only one intensional operator in STI, which is the propositional attitude verb, but under Keshet's (2011) split intensionality theory, the intensionality of the propositional attitude verb is split into two intensional operators: one is the propositional attitude verb and the other is the operator ^, which can be inserted lower than the propositional attitude verb. In this sense, the *de re* reading can be derived as long as the DP c-commands the operator ^.

According to Reinhart (1997), QR is not able to apply across finite clauses. For instance, in (7) the universal quantificational phrase *everybody* is embedded in the *that*-clause, and does not have the read where *everybody* takes the scope over *someone*.

(7)   Someone said that everybody was on time.

In this sense, one of the advantages for splitting the intensionality is that the DP does not have to move beyond the finite embedded clause at LF, which is shown in structure (5) with one finite embedded clause.
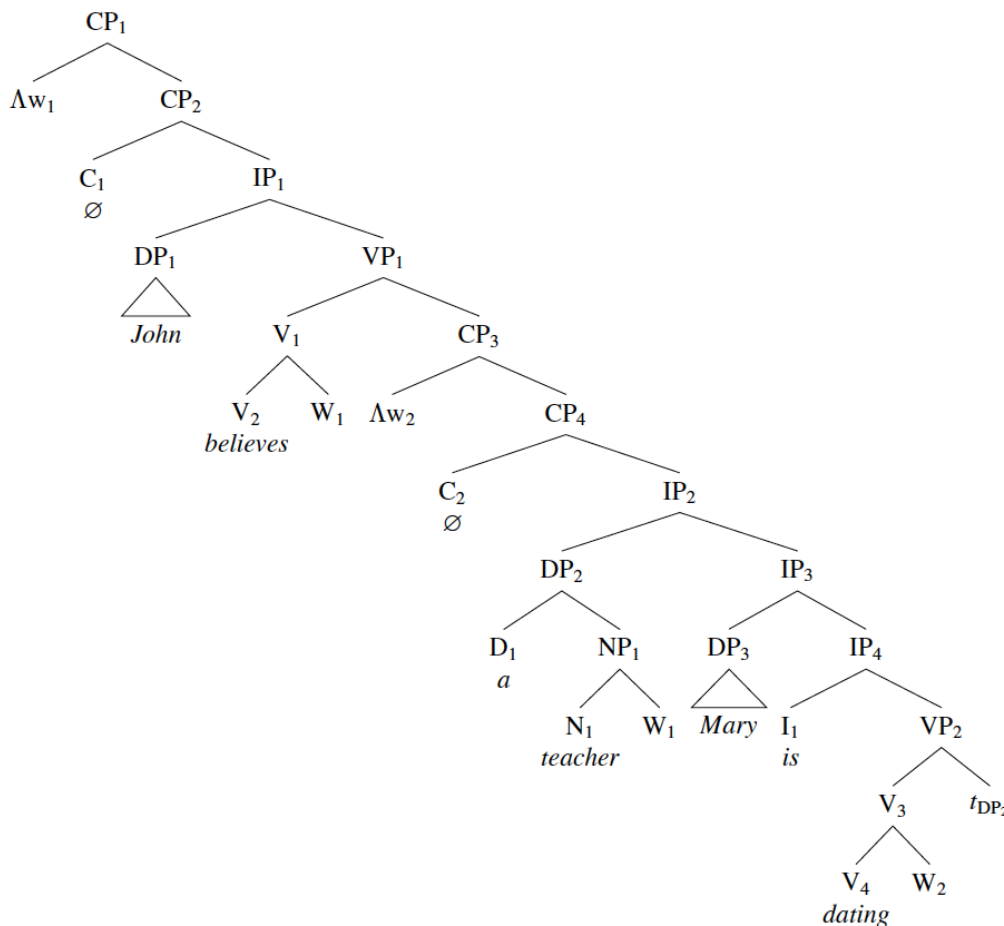
2.3. A system incorporating world variables

Both STI and the split intensionality theory explain the *de re/de dicto* ambiguity by analyzing the c-command relationship between the DP and the intensional operator: in STI, the intensional operator is the propositional attitude verb; under the split intensionality theory, the intensionality is split between the propositional attitude verb and the operator ^.

  Under the world variable account (see Cresswell 1990; Percus 2000; Fintel & Heim 2011), the *de re/de dicto* ambiguity is derived by interpreting the DP in the actual world or the attitude holder's belief worlds.

  Under this account, every N and V has a phonologically null pronoun W (i.e., a world pronoun) as its sister, which bears an index. For instance, in (8) *believes* and *teacher* have W_1 as its sister and this world pronoun W_1 is indexed as 1. In addition, every world pronoun is bound by an operator 'Λw_i'. It is interpreted by the following rule: $[[\Lambda w_i\ XP]]^w = [\lambda w_i: [[XP]]^{w(i \to w_i)} = T]$. This rule can be understood as: when the operator is combined with an XP, the world pronoun bound by this operator (i.e., the world pronoun that has the same index as the operator) in the XP is interpreted in the actual world if the operator is in the matrix clause, and interpreted in the attitude holder's belief worlds if the operator is in the embedded clause. For instance, in (8) the index "1" in W_1 indicates that it is bound by Λw_1. Since Λw_1 is in the matrix clause and combined with CP_2, W_1 which is the sister of *believes* and *teacher* is interpreted in the real world. Since Λw_2 is in the embedded clause and combined with CP_4, W_2 which is the sister of *dating* is interpreted in the attitude holder's belief worlds.

(8) Operators binding the world pronouns

```
                CP₁
            ┌────┴────┐
          Λw₁        CP₂
                  ┌────┴────┐
                 C₁        IP₁
                 Ø     ┌────┴────┐
                      DP₁        VP₁
                       △     ┌────┴────┐
                     John    V₁        CP₃
                          ┌───┼───┐  ┌──┴────┐
                         V₂  W₁  Λw₂         CP₄
                       believes        ┌──────┴──────┐
                                      C₂            IP₂
                                      Ø        ┌─────┴─────┐
                                             DP₂          IP₃
                                         ┌────┴────┐  ┌────┴────┐
                                        D₁        NP₁ DP₃       IP₄
                                         a      ┌──┴──┐ △   ┌────┴────┐
                                              N₁   W₁ Mary I₁       VP₂
                                            teacher          is  ┌────┴────┐
                                                                V₃       t_DP₂
                                                              ┌──┴──┐
                                                             V₄    W₂
                                                           dating
```

In this sense, if the world pronoun which is the sister of *teacher* is bound by the matrix operator Λw₁, *a teacher* is interpreted relative to the actual world and sentence (1) has a *de re* reading, which is shown in (8); if the world pronoun which is the sister of *teacher* is bound by the embedded operator Λw₂, *a teacher* is interpreted relative to the attitude holder's belief worlds and sentence (1) has a *de dicto* reading.

Under this account, the relative structural relationship between the propositional attitude verb and the DP is not relevant to the *de re/de dicto* ambiguity. QR can still be applied during the derivations, but the position that the DP can move to by QR is irrelevant.

2.4. Predictions

One major difference among the above three theories is how far the DPs should move at LF in order to get the *de re* reading: in the traditional scope theory of intensionality, the DP needs to move to a position that c-commands the verb; under Keshet's (2011) split intensionality theory, the DP does not need to move to a position that c-commands the verb but only to a position that c-commands the operator ^; and in the system incorporating world variables, the DP does not need to move to a certain position to get the *de re* interpretation.

Due to this major difference, one prediction would be that in the traditional scope theory of intensionality, the *de re* reading will not be possible if the verb is located outside an island containing the DP; under Keshet's (2011) split intensionality theory, the *de re* reading will not be possible if the operator ^ is outside an island where the DP is in; in a system incorporating world variables, the availability of the *de re* reading is not restricted by where the DP can move to.

In the following section, I use the data from Mandarin where the universal quantificational phrases can move overtly to show that in some situations even if the QP cannot move to a position that c-commands the propositional attitude verb or the operator ^, the *de re* reading is still available.

3.   The SOV order in Mandarin
3.1. A brief background

It is a historical debate whether Mandarin Chinese is an underlyingly SVO language (see Chu 1979; Huang 1982; Mei 1979; Sun & Givon 1985) or an underlyingly SOV language (see Tai 1973; Travis 1984; Li 1990; Gao 2008). In this paper, I assume that Mandarin Chinese is an underlying SVO language. I further assume that the SOV order is derived by the movement of the post-verbal object to the pre-verbal/post-subject position (see Gao 1994; Qu 1994; Shyu 1995; Ernst and Wang 1995; Li 1996, Tsai 2012)[4].

Example (9) shows the underlying SVO order, and example (10) shows the derived SOV order in Mandarin Chinese.

(9)   Lili   yijing   chi   wan   **wanfan** le
       Lili   already eat   finish dinner   PFV
       'Lili has already finished eating dinner.'

---

[4] Psycholinguistic evidence that supports the view that SOV in Mandarin is derived by movement can be found in Kuo (2009).

(10) Lili **wanfan_i** yijing chi wan t_i le
    Lili dinner already eat finish PFV
    'Lili has already finished eating dinner.'

Example (11) shows how the SOV order is derived from the SVO order when the object is a universal quantificational phrase.

(11) wo **mei-bu dianying_i** dou hen xihuan t_i
    1.SG every-CLF movie each/all very like
    'I like every movie very much.'

3.2. Syntactic islands in SOV order

I have shown in example (11) that in simple sentences, the universal quantificational phrase can occur in the pre-verbal object position. In examples (12a) and (12b), I show that the universal quantificational phrase occurs in a pre-verbal position in the matrix clause when it is originally the subject/object of the embedded predicate. In (13a), the universal quantificational phrase is in the embedded clause preceding the embedded adjective; while in (13b), the universal quantificational phrase moves to precede the matrix verb[5].

(12a) Steven juede [αP **mei-bu zhengzai shangying de jilu pian**
      Steven think every-CLF PROG on.show POSS documentary film
      dou hui hen kepa]
      each/all will very frightening
      'Steven thinks that every documentary film that is on will be very frightening.'

(12b) Steven **mei-bu zhengzai shangying de jilu pian_i** dou
      Steven every-CLF PROG on.show POSS documentary film each/all
      juede [αP t_i hui hen kepa]
      think will very frightening
      'Steven thinks that every documentary film that is on is very frightening.'

In (13a), the universal quantificational phrase is the object of the embedded predicate and precedes the embedded verb; while in (13b), it moves to precede the matrix verb.

(13a) ta rang [αP wo **mei-ben shu_i** dou mai t_i]
      3.SG make 1.SG every-CLF book each/all buy
      'S/he makes me buy every book.'

(13b) ta **mei-ben shu_i** dou rang [αP wo t'_i mai t_i]
      3.SG every-CLF book each/all make 1.SG buy
      'S/he makes me buy every book.'

---

[5] The syntactic and semantic properties of *dou* 'all/each' do not affect the judgment at all, so I leave aside the issues of *dou* 'each/all' throughout the discussion.

Examples (12a)-(13b) show that when the universal quantificational phrase is in an embedded clause, it can move to the pre-verbal position in the matrix clause, across the clause boundary. Due to the lack of tense morphology in Mandarin Chinese, it is not clear to me whether the embedded clause αP is finite or not, but it is clear that α is not a syntactic island for the movement of the universal quantificational phrase to the matrix clause.

However, sometimes the universal quantificational phrase cannot move across the embedded clause boundary to the position which precedes the matrix verb. This is shown by (14a)-(15b). In (14a), the universal quantificational phrase is the object of the embedded verb, and it moves to the position which precedes the embedded verb; however, it cannot occur in the position which precedes the matrix verb, which is shown in (14b). The same is true of (15a) and (15b).

(14a)   Bucky renwei [βP Steven **mei-bu   zhengzai   shangying de   jilu**
        Bucky believe   Steven every-CLF PROG   on.show   POSS   documentary
        pian$_i$ dou   kan   guo le   t$_i$]
        film   each/all watch PFV PFV
        'Bucky believes that Steven has watched every documentary film that is on.'

(14b) *Bucky **mei-bu   zhengzai   shangying de   jilu   pian**$_i$ dou
        Bucky every-CLF PROG   on. show   POSS documentary film   each/all
        renwei [βP Steven **t'**$_i$ kan   guo le   t$_i$]
        believe   Steven   watch PFV PFV
        Intended: 'Bucky believes that Steven has watched every documentary film that is on.'

(15a)   Zhangsan zhidao [βP wo **mei-chang bisai**$_i$ dou   kan   guo le   t$_i$]
        Zhangsan know   1.SG every-CLF match   each/all watch PFV PFV
        'Zhangsan knows that I have watched every match.'

(15b) *Zhangsan **mei-chang bisai**   dou   zhidao [βP wo **t'**$_i$ kan   guo le   t$_i$]
        Zhangsan every-PFV match   each/all know   1.SG   watch PFV PFV
        Intended: 'Zhangsan knows that I have watched every match.'

Assuming that SOV order is derived by movement, it is highly likely that the reason why in (14b) and (15b) the universal quantificational phrase cannot occur in the pre-verbal position in the matrix clause is that the embedded clause βP is a syntactic island which prevents the universal quantificational phrase from moving to the matrix clause.

In short, the universal quantificational phrase cannot always move out of an embedded clause to a position preceding the matrix verb.

Recall that in the very beginning, I adopt the assumption that the SOV order is derived by movement. I have not shown direct evidence to prove that SOV is derived by movement, but considering (12a)-(15b), SOV being base-generated cannot explain why SOV is possible in (12b) and (13b), but not possible in (14b) and (15b): if a DP can be base-generated in a pre-verbal position, one has to assume that SOV order is possible depending on which verb the sentence has. This means that SOV is possible for the verb in (12b) and (13b), but not the verb in (14b) and (15b). This is an odd assumption.

4. Different approaches to the Mandarin data
4.1. The traditional scope theory for intensionality

In Section 3, I show that in example (14a), the embedded clause is a syntactic island for the universal quantificational phrase preventing it from moving to the matrix clause. Assuming that when α is an island for a particular overt movement in narrow syntax, αP is also an island for the corresponding covert movement at LF[6], the universal quantificational phrase in (14a) cannot move to a position that precedes the matrix verb at LF, either. In this sense, in (14a), by applying QR, the farthest position that the universal quantificational phrase can move to is the edge of the embedded clause, which is always below the propositional attitude verb. This is shown in (16).

(16) LF for example 14(a)
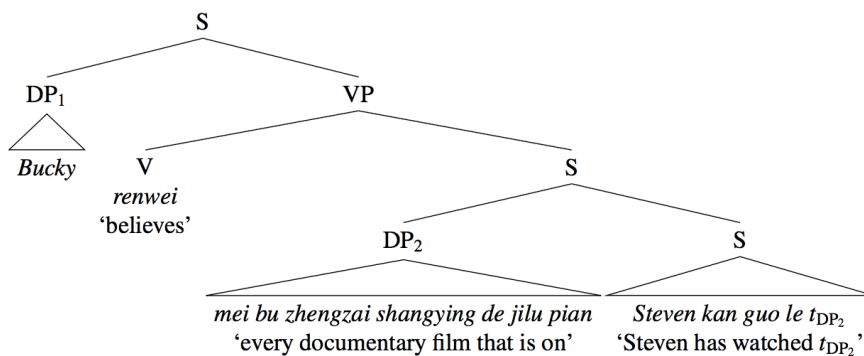
Since the universal quantificational phrase is always below the propositional attitude verb, the traditional scope theory for intensionality predicts that (14a) can only receive the *de dicto* reading, and cannot receive the *de re* reading, which requires the universal quantificational phrase c-commands the propositional attitude verb.

Under the traditional scope theory for intensionality, in order for (14a) to receive a *de re* reading (14b) should have been possible. Thus, the fact that (14b) is ungrammatical suggests that the traditional scope theory for intensionality cannot explain why (14a) can receive a *de re* reading. (14a) is repeated as (17) below.

(17) Bucky   renwei   [βP Steven **mei-bu**   **zhengzai**   **shangying de**   **jilu**
      Bucky   believe   Steven   every-CLF   PROG   on.show   POSS   documentary
      piani dou   kan   guo   le   ti]
      film   each/all watch   PFV   PFV
      'Bucky believes that Steven has watched every documentary film that is on.'

One imaginary situation for the *de re* reading is the following.

In the cinema, there are six movies on show. Three of them are documentary films and three of them are comedies. Steven points to those three documentary films, and tells Bucky that he has watched all of them. However, by mistake, Bucky believes those three

---

[6] This assumption is compatible with the idea that the distinction between 'overt' and 'covert' movement is whether the head or the tail of the chain is pronounced (see Bobaljik 1995 and many others): no matter whether the universal quantification phrase in Mandarin moves 'covertly' (i.e., the tail of the chain gets pronounced) or 'overtly' (i.e., the head of the chain gets pronounced), it moves along the same path, and obeys the same island constraints.

documentary films to be horror movies. Thus, one can say (17), although in Bucky's mind, those three movies are horror movies. In other words, in (17), the universal quantificational phrase is interpreted relative to the actual world (i.e., in the actual world, those three movies are indeed documentary films), but not interpreted relative to Bucky's belief world (i.e., what Bucky believes in his mind).

4.2. Keshet's (2011) account

Under Keshet's (2011) split intensionality theory, the LF for the *de re* reading of example (17) is shown in (18).

(18) LF for the *de re* reading of example (17)
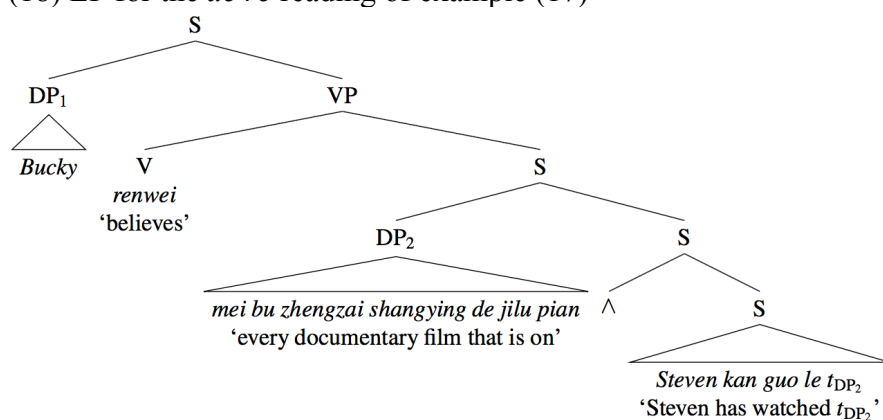
Since the operator ^ can be inserted freely, it can be inserted as the sister to the embedded clause. Then, by applying QR, the universal quantificational phrase can further move to the edge of the embedded clause, as long as it does not move across the matrix verb *renwei* 'to believe'. Any region c-commanding the matrix verb is outside the clause boundary. In (18), both the universal quantificational phrase and the operator ^ are at the edge of the embedded clause and thus are within the clause boundary, but the universal quantificational phrase is at the farther edge. Therefore, the *de re* reading of (17) can be explained under Keshet's (2011) split intensionality theory.

   Since the intensionality is split between the matrix verb and the operator ^, the universal quantificational phrase does not have to c-command the matrix verb to receive the *de re* interpretation. In (18), although the universal quantificational phrase is below the matrix verb, but it is above the operator ^, and thus can receive the *de re* interpretation.

   Keshet's (2011) split intensionality theory enables (17) to get the *de re* reading by inserting the operator ^ within the embedded clause so that the universal quantificational phrase does not need to move out of the embedded clause boundary. However, sentences like (19) [7], where the highest operator ^ is located outside the clause boundary, cannot be explained under Keshet's (2011) split intensionality theory.

---

[7] Many thanks to Seth Cable for providing the English example in (18). He also points out that *De Re* reading is available in (18) in English.

(19) ruguo [TP Mary renwei [ΦP **mei-ge jiaoshou** dou shi xuesheng]],
     if   Mary   believe   every-CLF professor each/all be   student
     ta   hui   renwei   Bill shi ge   xuesheng
     3.SG would believe   Bill be   CLF student
     'If Mary believed every professor was a student, she would believe Bill is a student.'

Like sentence (17), the clause Φ that the verb *renwei* 'to believe' takes as its complement is also an island for the universal quantificational phrase and it cannot move to a position that precedes the verb *renwei* 'to believe'. This is shown in (20).

(20)   *ruguo Mary **mei-ge jiaoshou**ᵢ dou renwei [ΦP tᵢ shi xuesheng],
      if   Mary every-CLF professor each/all believe   be   student
      ta   hui   renwei   Bill shi ge   xuesheng
      3.SG would believe   Bill be   CLF student
      Intended: 'If Mary believed every professor was a student, she would believe Bill is a student.'

Under Keshet's (2011) split intensionality theory, the LF for (19) is shown in (21).

(21) LF for example (19)
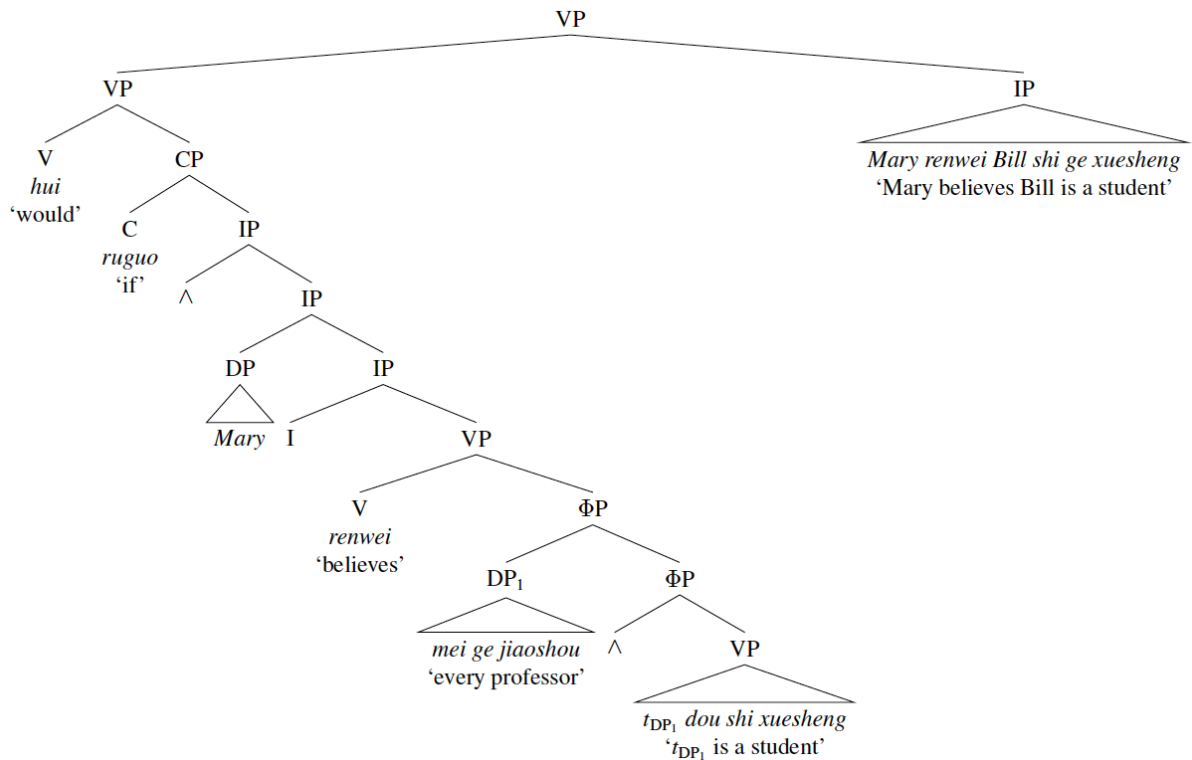
In (21), there are two sources for intensionality: the modal verb *hui* 'would' and *renwei* 'to believe'. The operator ^ is split from the modal verb and inserted at the edge of the IP, and another operator ^ is split from the verb *renwei* 'to believe' and inserted at the edge of ΦP. In order to get the *de re* reading for (19), the universal quantificational phrase has to c-command the highest operator ^, which is split from the modal *hui* 'would'. Recall that under the split

intensionality theory, the *de dicto* region is below the operator ^, and in this case below the highest operator ^. However, I have shown that ΦP is an island for the universal quantificational phrase, and thus it cannot move outside ΦP; the farthest position it can move to is the edge of ΦP. Thus, in (21), the universal quantificational phrase is not able to c-command the highest operator ^, and thus it cannot receive a *de re* reading. Thus, the *de re* reading of (19) is left unexplained under Keshet's (2011) split intensionality theory.

An imaginary situation for the *de re* reading of sentence (19) is the following.

In the classroom, there are students and three professors. Mary just arrives and does not know there are three professors and thinks everyone is a student by mistake. Bill is one of the three professors. In this case, one can say sentence (19), when in Mary's mind, every one in the classroom is a student. Thus, *mei ge jiaoshou* 'every professor' is interpreted relative to the actual world, not to Mary's belief worlds.

4.3. World variables account

In this section, I show that under the world variables account (see Cresswell 1990; Percus 2000; Fintel & Heim 2011), the *de re* reading of sentences (17) and (19), which are repeated as (22) and (23) below, can both be explained.

(22) Bucky   renwei   [βP Steven   **mei-bu**   **zhengzai**   **shangying de**   **jilu**
      Bucky   believe   Steven   every-CLF   PROG   on.show   POSS   documentary
      pian$_i$ dou   kan   guo   le   t$_i$]
      film   each/all   watch   PFV   PFV
      'Bucky believes that Steven has watched every documentary film that is on.'

(23) ruguo   [TP Mary   renwei   [ΦP **mei-ge**   **jiaoshou**   dou   shi   xuesheng]],
      if   Mary   believe   every-CLF   professor   each/all   be   student
      ta   hui   renwei   Bill   shi   ge   xuesheng
      3.SG   would   believe   Bill   be   CLF   student
      'If Mary believed every professor was a student, she would believe Bill is a student.'

Recall that under the traditional scope theory for intensionality, in order to get the *de re* reading, the universal quantificational phrase should move to a position that c-commands the verb *renwei* 'to believe' in (22), and to a position that c-commands the modal *hui* 'would' in (23). However, βP and ΦP are syntactic islands that prevent the universal quantificational phrase from moving out. Thus, the traditional scope theory for intensionality cannot explain why (22) and (23) can have a *de re* reading.

Under Keshet (2011)'s split intensionality theory, for (22) to get the *de re* reading, the universal quantificational phrase does not have to move to a position that c-commands the verb *renwei* 'to believe'; instead, it only needs to move to a position that c-commands the operator ^, which is split from the verb *ren wei* 'to believe' and within βP. Thus, the *de re* reading of (22) can be explained by moving the universal quantificational phrase to the edge of βP. However, for (23) to get the *de re* reading, the universal quantificational phrase must move out of ΦP because it needs to c-command both the lower and highest operator ^, and the highest operator ^, which is split from the modal *hui* 'would' is outside ΦP. Thus, the *de re* reading of (22) can be explained, but the *de re* reading of (23) cannot be explained under Keshet's (2011) split intensionality theory.

However, under the world variables account, the *de re* interpretation is no longer derived by the relative c-command relationship between the DP and the intensional operator, but by the binding relationship between the operator 'Λw$_i$' and the world pronoun W$_i$. Thus, whether the DP can receive a *de re* interpretation is independent from where the DP can move by applying QR. In other words, the syntactic islands are not relevant under the world variables account. This is shown by (24) and (25), which are the LF representations of (22) and (23).

(24) **Λw₁** Bucky renwei-**w₁** [βP**Λw₂** Steven mei-bu zhengzai shangying de
　　　 Bucky believe 　　　Steven every-CLF PROG 　　on.show POSS
　　 jilu pian$_i$-**w₁** dou kan-**w₂** guo le t$_i$
　　 documentary film each/all watch PFV PFV
　　 'Bucky believes that Steven has watched every documentary film that is on.'

(25) **Λw₁** hui **Λw₂** ruguo [TPMary renwei-**w₂** [ΦP *Λw₃* mei-ge
　　　 would 　　 if 　　Mary believe 　　　every-CLF
　　 jiaoshou-**w₁** dou shi xuesheng-*w₃*]], ta renwei Bill shi ge
　　 professor each/all be student, 3.SG believe Bill be CLF
　　 xuesheng
　　 student
　　 'If Mary believed every professor was a student, she would believe Bill is a student.'

In (24) and (25), both the N *jilu pian* 'documentary film' and the N *jiaoshou* 'professor' have W₁ as its sister, which is bound by the highest Λw₁. Thus, the N *jilu pian* 'documentary film' and *jiaoshou* 'professor' are both interpreted relative to the actual world, and receives the *de re* reading.

5. Problem and Conclusion

In this paper, I adopt the assumption that at LF, the farthest position the universal quantificational phrase can move to is the edge of the embedded clause when the embedded clause is a syntactic island. However, in Mandarin, the universal quantificational phrase can also occur in pre-subject position. For (14b) and (15b), although SOV is not allowed, but OSV is possible, which is shown in (26) and (27).

(26) **mei-bu zhengzai shangying de jilu pian$_i$** Bucky dou
　　 every-CLF PROG on.show POSS documentary film Bucky each/all
　　 renwei [βP Steven t'$_i$ kan guo le t$_i$]
　　 believe 　　 Steven watch PFV PFV
　　 'Bucky thinks that Steven has watched every documentary film that is on.'
(27) **mei-chang bisai** Zhangsan dou zhidao [βP wo t'$_i$ kan guo le t$_i$]
　　 every-CLF match Zhangsan each/all know 　　1.SG watch PFV PFV
　　 'Zhangsan knows that I have watched every match.'

However, it is odd to assume that the universal quantificational phrase in (14a) and (15a) cannot move to a pre-verbal/post-subject position in the matrix clause, but can move directly to the pre-subject position in the matrix clause, while the universal quantificational phrase in (12a) and (13a) can move to both the pre-verbal/post-subject and pre-subject position in the

matrix clause. It is possible that the QP in the pre-subject position in the matrix clause is base-generated, and not derived by movement. In this sense, the embedded QP cannot move to the pre-subject position at LF, and can only move to as far as the pre-verbal position in the matrix clause.

In conclusion, I show in this paper that the traditional scope theory for intensionality and Keshet's (2011) split intensionality theory cannot fully explain the data from Mandarin; while the data can be explained by the world variables account.

Acknowledgments

Rong Yin
University of Massachusetts Amherst
ryin@umass.edu

References

Bobaljik, J. (1995). Morphosyntax: the Syntax of Verbal Inflection. PhD thesis. MIT.
Cable, S. (2016). The *De Re/De Dicto* Ambiguity. Unpublished lecture notes. UMass Amherst.
Chu, C.C. (1979). Hàn-yǔ de cǐ-xū yǔ cǐ-xū bián-qian zhōng de wēn-tīái (Problems of word order and word order change in Chinese). Yǔ-yán-xué lùn-jí: Li-lùn, yìng-yóng jì hàn yǔ-fǎ. Crane Publishing Co, Taipei. [in Chinese]Cresswell, M. J. (1990). *Entities and Indices*. Kluwer, Dordrecht.
Ernst, T. & C. Wang (1995). Object preposing in Mandarin Chinese. *Journal of East Asian Linguistics* 4:3, pp. 235-260.
Fodor, J. (1970). *The linguistic description of opaque contents*. PhD thesis. MIT.
Gao, Q. (1994). Focus Criterion: Evidence from Chinese. Paper presented at the 6th North American Conference on Chinese Linguistics. Los Angeles: University of Southern California.
Gao, Q. (2000). Argument Structure, HPSG and Chinese Grammar. PhD thesis. Ohio State University.
Heim, I & K. von Fintel. (2011). *Intensional Semantics*. [Ms]. MIT.
Huang, C.-T. J. (1982). Logical Relations in Chinese and the Theory of Grammar. PhD thesis. MIT.
Keshet, E. (2011). Split intensionality: a new scope theory of de re and de dicto. *Linguistics and Philosophy* 33:4, pp. 251–283.
Kuo, Grace. (2009). Mandarin Non-Canonical Word Orders in Sentence Processing. Paper presented at the 2nd International Conference on East Asian Linguistics (ICEAL-2). Vancouver, Canada.
Ladusaw, W. (1977). Some problems with tense in PTQ. *Proceedings of the 6th Texas Linguistic Forum*, pp. 89-102.
Li, J. J. (1996). Preverbal NP Positions in Mandarin Chinese. PhD thesis. Tucson: The University of Arizona.
Li, A. (1990). Order and Constituency in Mandarin Chinese. Kluwer, Dordrecht.
Mei, K. (1979). Is modern Chinese really a SOV language. Tang C. and I. Li (eds.), *Proceedings of the 1979 Asian and Pacific Conference on Linguistics and Language Teaching*, Taipei, Student Books, pp. 275–297.
Montague, R. (1973). The proper treatment of quantification in ordinary English. Hintikka, K. J. J., J. M. E. Moravcsik & P. Suppes (eds.), *Approaches to Natural Language*, Reidel/Dordrecht, pp. 221–242.
Ogihara, T. (1992). Temporal reference in English and Japanese. PhD thesis. University of Texas at Austin.
Ogihara, T. (1996). *Tense, Attitudes, and Scope*. Kluwer, Dordrecht.
Percus, O. (2000). Constraints on some other variables in syntax. *Natural Language Semantics* 8:3, pp. 173-229.

Qu, Y. (1994). Object Noun Phrase Dislocation in Mandarin Chinese. PhD thesis. Vancouver: University of British Columbia.

Reinhart, T (1997). Quantifier-Scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:4, pp. 335-397.

Russell, B. (1905). On denoting. *Mind* 14:56, pp. 479-493.

Shyu, S. (1995). The Syntax of Focus and Topic. PhD thesis. Los Angeles: University of Southern California.

Stowell, T. (1993). Syntax of tense. [Ms]. UCLA.

Sun, C. & T. Givón. (1985). On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language* 61:2, pp. 329-351.

Tai, H. J. (1973). Chinese as an SOV Language. Corum, C., J. C. Smith & A. Weister (eds.), *Proceedings of the 9th Chicago Linguistic Society*, Chicago, CLIS, pp. 659-671.

Travis, L. (1984). Parameters and Effects of Word Order Variation. PhD thesis. MIT.

Standards of comparison and the case of Spanish "*que - de* alternation"

Laura Vela-Plo

This paper offers a formal syntactic and semantic analysis of standards of comparison in Spanish, in addition to presenting a parametric analysis that accounts for the variation in standards of comparison cross-linguistically. Two independent parameters are argued to account for the properties of standards of comparison in languages typologically as diverse as Spanish, English, Japanese or Turkish: (i) a syntactic parameter, by which either phrasal, or both phrasal and clausal structures are allowed in the complement position of the standard marker; and (ii) a revised semantic parameter that restricts the semantic type of the standard.

1. Introduction

The syntactic and semantic properties of comparative structures are still a topic of great debate. In order to examine comparative structures as a whole, we need a deeper understanding of the individual pieces that compose the puzzle that comparatives represent. For this reason, the focus of this paper lies on one of the main components of these structures: the standard of comparison.

The standard of comparison sets the reference from which the target of comparison is described. In a simple sentence like (1), for instance, Ana is the target of the comparison; she is being compared to Jon (the standard of comparison) with respect to their individual levels of intelligence, and Jon's level of intelligence has been set as the referent for the comparison.

(1)   Ana is smarter than Jon.

How is it that we come to interpret a standard of comparison such as 'Jon' as Jon's maximum level of intelligence? And more importantly for our concerns here, how can we derive a compositional analysis that accounts for the syntactic and semantic properties of standards of comparison? What are the constraints that apply to standards of comparison cross-linguistically? These are the main concerns that this paper sets out to address. By answering these questions, we will be able to determine the limits of the cross-linguistic variation in standards of comparison, as well as to lay down the first piece of the puzzle on comparative structures.

The main claim endorsed in this paper is that two independent parameters[1] are necessary to account for the properties of standards of comparison cross-linguistically: a syntactic one, namely the licensing of a [-/+ CLAUSAL] structure, as previously proposed in the literature; and a revised semantic parameter, which lies in the possibility to license [-/+ DEGREE] denoting standards.

In Sections 1.1 and 1.2 I will review some of the most recent literature and analyses on the syntactic and semantic points of variation in standards of comparison. In Section 2, I will examine the syntactic and semantic properties of standards of comparison in a specific Romance language, Spanish. This language is especially interesting for the purposes of this research because the types of standards of comparison present in this language *prima facie* seem to be compatible with the syntactic and semantic variation patterns that have been proposed so far. However, we will see that the semantic analyses that we can find in the literature cannot account for the whole comparative paradigm in this language and that some refinements are necessary. Hence, in Section 3, I will provide a novel semantic description and classification of standards of comparison in Spanish (which will be classified in three main types; Sections 3.1, 3.2 and 3.3). Finally, in Section 3.4 I will show how the analysis of the properties of standards of comparison in Spanish fits in with the cross-linguistic classification and parametric description of these phenomena.

An additional outcome of the analysis of Spanish inequality comparatives and the '*que-de* alternation' examined in this paper is that, at least in Spanish, the sense of comparison in a comparative structure is contingent on both the presence of a comparative marker (*más* 'more') and the choice of standard marker (*que* 'that' or *de* 'of, from').

1.1. Previous syntactic analyses

The syntactic properties of standards of comparison are subject to important cross-linguistic variation.

On the one hand, some languages only allow phrasal arguments in the complement position of the standard marker. I will be using the label [- CLAUSAL] for those standards of comparison that do not seem to be propositional. Languages such as Turkish (Hofstetter 2009, see (2)), Japanese (Beck et al. 2004, Sudo 2015), Mandarin (Fu 1978; Xiang 2003), or Hindi-Urdu (Bhatt & Takahashi 2008), for example, do not allow this type of clausal, CP arguments in the complement position of their standard marker unless they are embedded under a nominal in a relative clause (RC)[2] or nominalized, as in the example in (2c).

(2) a. Maria Peter'den hızlı koştu.   (Turkish; Hofstetter 2009:189-190)
         Maria Peter.ABL fast   run.PST.3SG
         'Maria ran faster than Peter.'

---

[1] Througout this paper, parameters are regarded as feature specifications on lexical items that vary within languages and derivations, arguably on the standard marker that introduces the standard of comparison, rather than parameters which globally define a language in the sense of the Principles and Parameters framework (cf. Eguren, Soriano & Mendikoetxea 2016, for a recent review on this topic). Thus, parameters are not considered binary switches that determine whether a particular language allows a syntactic structure or not. As will be seen later in the descriptions of the languages examined in this paper, and as it is summarized in Tables 2 & 3, both parametric options might be active within a language.

[2] I will be assuming that a RC under a nominal head is still a phrasal argument (a nominal with a clausal modifier).

    b.   *Maria Peter'den (hızlı) koştu   hızlı koştu.
             Maria Peter.ABL (fast) run.PST.3SG fast   run.PST.3SG
             Intended: 'Maria ran faster than Peter ran.'
    c.   Maria benim düşündüğümden   zengin.
             Maria my   think.PTCP.1SG.ABL rich
             'Maria is richer than I thought.'

On the other hand, some languages allow their standard markers to take both phrasal and clausal complements [-/+ CLAUSAL]. These are languages such as English (Heim 1985; Kennedy 1999, a.o.), Spanish (Brucart 2003), Italian, Hungarian and Russian (Bacskai-Atkari 2014). The sentence in (3) shows an example of a phrasal standard in English, in which clausal expansion is not permitted. In this case, the complement of the standard marker *than* is a Measure Phrase (MP) of a [- CLAUSAL] nature.

(3)   This mountain is higher than 1200 meters (*are/*is). / MP   [- CLAUSAL]

In contrast with (3), the example in (4) licenses a standard that allows clausal expansion: *than Jon is*. The reduced version of (4), *than Jon,* could be thus analysed either as having an underlying clausal structure to which a deletion process has applied (a *reduction* analysis; cf. (5)); or as involving a phrasal standard (a *direct* analysis; cf. (5b)). Within reductionist analyses, the underlying structure of (4a) has been represented as in (5a) since Chomsky (1977), with movement of a degree operator (similar to a *wh*-movement) to the left periphery of the clause and deletion of all redundant material, the gradable predicate in this case (also known as *comparative deletion*).

(4)   a.   Ana is smarter than Jon (is).
       b.   Ana is smarter than him (*is).
(5)   a.   …than [$_{CP}$ *Op$_i$* Jon is *d$_i$* ~~smart~~]   [+ CLAUSAL]
       b.   …than [$_{DP}$ him][3]   [- CLAUSAL]

The sentence in (4b), in contrast, is apparently phrasal, since the standard of comparison does not surface wth the case mark that corresponds to the subject of an inflected clause, as could be expected under the deletion analysis. Furthermore, the alternative with clausal expansion leads to unacceptability. Hence, a *direct* analysis, namely a *what-you-see-is-what-you-get* interpretation, is available, which could potentially be represented as in (5b).

    Moreover, the example in (6) is a comparative sentence with a clearly clausal standard in English, also called *subcomparative*. It is common practice to analyse these cases as also involving degree operator movement to the left periphery of its clause (5b), just as in (4a).

(6)   a.   This table is longer than that door is (wide).
       b.   This table is longer than [$_{CP}$ *Op$_i$* that door is *d$_i$* wide]   [+ CLAUSAL]

---

[3] Hankamer (1973), Napoli (1983) and Kennedy (1999), among others, argue for a direct analysis as represented in (5b). Even though a direct analysis would be the most straightforward analysis, it might not be the only analysis for (4b). See Pancheva (2006) for an alternative analysis involving a small clause. For the necessary application of the direct analysis in comparatives in other languages, such asHindi-Urdu, see Bhatt & Takahashi (2007, 2011).

In sum, the first difference that we observe with respect to the syntactic properties of the standard of comparison has to do with the nature of the complement of the standard, and in particular, with whether both phrasal and clausal complements are licensed in this position, or whether only phrasal complements are. We could translate this point of variation observed within and across languages as a [-/+CLAUSAL] parameter.

1.2. Previous semantic analyses

The standard of comparison has traditionally been represented semantically as a type of definite description that denotes a maximal degree (von Stechow 1984, Rullmann 1995). Hence, within the traditional analysis of standards of comparison it is common to assume that comparatives with a clausal standard such as (6a) above are degree-denoting (see e.g. Bresnan 1973; von Stechow 1984; Bhatt & Pancheva 2004). This degree is mapped onto a scale, an abstract representation that can be used to measure and compare objects with respect to a particular gradable property (Kennedy 2006, Bale 2008).

Nonetheless, some authors have argued that in addition to degree-denoting standards (type *d*), there are also cases of comparisons of individuals in which the standard of comparison exhibits a referential definite that denotes an individual (Hoeksema 1983; Heim 1985; Kennedy 1999, 2007; Bhatt & Takahashi 2007, 2011), represented by the semantic type *e*. For instance, Kennedy (2007) argues that Japanese has only individual comparison, and thus, that all standards of comparison in this language are DPs (either simple DPs that are individual-denoting as in (6a) or complex DPs modified by a RC that are degree-denoting) that denote individuals. Shimoyama (2012:85, fn 3) argued in favour of the weaker claim that says that both degree and individual-denoting standards are available in Japanese by showing that the Japanese standard marker *yori* can take the name of a degree (a Measure Phrase) as its complement (see (7b)).

(7)   *Japanese*
    a.   Taroo-wa *Hanako* yori takusan(-no) hon-o   katta.
           Taroo-TOP Hanako than many(-GEN)   book-ACC bought
           'Taroo bought more books than Hanako.'
    b.   Kono mati-de-wa *gozyuu-meetoru*-yori takai biru-wa   kinsisareteiru.
           this   town-in-TOP fifty-meters-than   high building-TOP prohibited
           In this town, buildings higher than 50 meters are prohibited.

This proposal has been supported by Bhatt & Takahashi (2007, 2011). These authors also offer several arguments supporting the view that both types of standards are also available in Hindi-Urdu: degree-denoting ones such as 'three' in sentence (8b) and individual-denoting standards such as 'Bill' in (8a).[4]

(8)   *Hindi-Urdu*
    a.   John *Bill*-se   zyaadaa lambaa hai.
           John Bill-than more   tall   is
           'John is taller than Bill.'

---

[4] The reader should refer to the original papers for evidence supporting the argument that cases such as (8a) cannot be analysed as degree-denoting standards in disguise by appealing to a reduced clause analysis.

b.　Atif-ne　*tiin*-se　　zyaadaa kitaabẽ khariid-ii thĩ:
　　Atif-ERG three-than more　　books.F buy-PFV.f be.PST.FPl.
　　'Atif bought more than three books.'

In order to account for the Japanese constrast, Kennedy (2007) proposes a degree versus individual-denoting standard analysis (adaptation of Kennedy 2007, based on the work by Beck et al. 2004):

(9)　a.　⟦more $_{degree}$⟧ $= \lambda G_{<e,d>} \lambda d_d \lambda x_e [G(x) > d]$　　　　　　(7b)
　　b.　⟦more $_{individual}$⟧ $= \lambda G_{<e,d>} \lambda y_e \lambda x_d [G(x) > G(y)]$　　　(7a)

Sentence (7b) would be analysed as in (9a), with a standard that denotes a degree directly; whereas the representation in (9b) would be applied to individual standards such as the one in (7a). These types of standards express orderings between arbitrary individuals, and the degree of the standard is derived by applying the meaning of the gradable predicate (the base of comparison) to the two individuals that are being compared. Eventually, standards are of type *d* in both cases.

Bhatt & Takahashi (2007, 2011) follow a similar strategy to account for the variation observed in Hindi-Urdu standards of comparison (cf. (8)) and also posit two '*more*'s: one taking standards of type *d* and the other taking type *e* standards (*2-place and 3-place 'more'*).

Summarizing, the second attested point of variation in standards of comparison has to do with the denotation of the complement of the standard marker, and in particular, with the possibility of having degree and individual-denoting standards. We could translate this point of variation as a [DEGREE]/[INDIVIDUAL] semantic parameter.

In the next section, I will examine inequality comparatives in Spanish and introduce the *que-de* alternation that will be discussed in this paper. Building on previous analyses of Spanish comparatives and on some novel observations, I will show that the semantic parameter proposed in the literature is insufficient to account for the whole comparative paradigm in Spanish. On this basis, I will thus propose some refinements to this parameter and a new classification of Spanish comparatives. Finally, I will offer a compositional analysis of inequality comparatives in Spanish and a parametric classification that accounts for the cross-linguistic variation on standards of comparison in Section 3.

2. Spanish "que-de *alternation" and some novel observations*

Comparative structures in Spanish are especially interesting due to the variation that we find in this language. Specifically, there are two possibilities available in Spanish comparatives to introduce the referent of the comparison, namely the standard: the complementizer or conjunction *que* ('that') or the preposition *de* ('of, from').

According to the literature, *de* standards confront comparable "magnitudes" (either degrees or quantities), and have an obligatorily quantitative nature (Martínez 1987, Gutiérrez Ordóñez 1994, Romero Cambrón 1998). In contrast, *que* standards do not have a quantificational base; rather, they primarily represent individuals or properties (Brucart 2003, 2009). This contrast is illustrated in (10) and (11).

(10) a.   Mide   más {* que   /de } dos metros.
            measure more   than$_{que}$/than$_{de}$ two meters
            'It measures more than$_{de}$ two meters.'

      b.   Mide   más {que   /*de   } Ana.
            measure more than$_{que}$/than$_{de}$ Ana
            'It measures more than$_{que}$ Ana'.

(11) a.   Jon comió más que un dinosaurio.[5]
            Jon ate   more than a dinosaur
            'Jon ate more than$_{que}$ a dinosaur (would eat).'

      b.   Jon comió más de   un dinosaurio.
            Jon ate   more than a dinosaur
            'Jon ate more than$_{de}$ one dinosaur.' (maybe 2 or 3 dinosaurs)

In (10a), we observe that when the standard of comparison denotes a degree, as in the MP 'two meters', a *de* standard is the only acceptable option; while in (10b), the individual denoting standard *Ana* has to be introduced by *que*.[6]

    The minimal pair in (11a)-(11b) illustrates a similar contrast. In (11), I have given an example which illustrates the different roles of *que* and *de* in the standard of comparison and which is particularly interesting due to the ambiguous nature of *un* between the indefinite 'a' and the numeral 'one'. In the comparative with the *que* standard, the quantity eaten by two individual referents is compared ((11a), *Figure 1*), and *un dinosaurio* is interpreted as an indefinite referent, 'a dinosaur'; whereas in the case of the standard with *de* the article *un* has to be translated as the numeral 'one', 'than one dinosaur', alluding to the quantity of eaten elements ((11b), *Figure 2*).

Figure 1: Interpretation of (11).   *Figure 2: Interpretation of (11).*

---

[5] This sentence could also have the meaning 'Jon ate more than$_{que}$ a dinosaur.' (something qualitatively different, 'other than a dinosaur'), if *more* were focalized. That is, both a subject or object interpretation of the indefinite referent 'a dinosaur' are possible in (10a).

[6] Whether we need a direct or a reductionist analysis of the standard in sentences (9b) and (10b) is not a crucial distinction at this point. For now, we are interested in the distributon of *que* and *de* standards. The first observation is that *que* is not allowed with degree-denoting standards, but *de* is. The second observation is that the standard marker *de* cannot directly take what in the surface looks as an individual-denoting element, while this is allowed with *que*.

In what follows, I will focus on inequality comparatives that present standards that are obviously clausal (for example, (19) below) or comparatives with phrasal standards for which there is enough evidence for a direct phrasal analysis (as in sentences (12) and (13), for instance).

Moreover, *de* standards cannot take clausal complements unless they are embedded in a relative clause without antecedent, as shown by the contrast in (12a)-(12b) (in fact, *de* standards are known as 'relative comparatives' in traditional grammars of Spanish; Gutiérrez Ordóñez 1994, Romero Cambrón 1998, Sáez 1999, Brucart 2003, 2009, a.o.). No such restriction is found in *que* standards (see (13); cf. Gutiérrez Ordóñez 1994, on the possibilities when having a standard with *que* in Spanish).

(12) a.   *Ana ha conseguido más libros para el colegio de Jon necesita.
            Ana has obtained   more books for   the school   than Jon needs
            Intended: 'Ana got more books for school than_de Jon needs.'
     b.   Ana ha conseguido más libros para el colegio de los que Jon necesita.
            Ana has obtained   more books for the school   than the.PL that Jon needs
            'Ana got more books for school than_de Jon needs.'

(13) Ana ha conseguido más libros para el colegio que trabajos ha pedido su
     Ana has obtained   more books for   the school than essays   has asked her
     profesor.
     teacher
     'Ana got more books for school than_que her teacher ordered essays.'

Taking previous analyses into consideration, we would expect the pattern in (14).

(14) Potential analysis of Spanish standards of comparison (to be revised):
     A.   *de* standards: [DEGREE]   [- CLAUSAL]
     B.   *que* standards: [INDIVIDUAL] [- CLAUSAL] or [DEGREE] [+ CLAUSAL]

However, as I will show now, the semantic characterization of comparatives proposed in Section 1.2. is insufficient to capture the properties of Spanish comparatives: the picture is more complex than what the traditional analyses predict. This is so because we find cases that do not conform to the semantic [INDIVIDUAL]/[DEGREE] distinction, since (i) they do not involve a directly degree-denoting standard (via a MP as in (3) or degree abstraction as in (5a) or (6)), and further (ii) they are not individual-denoting standards (*e* type, as in (4b)) either.

2.2. More N than N comparatives

Consider the examples in (15) and (16). Sáez (1992) suggests that standards of comparison such as *que hombres* 'than men' in (15) could be analysed as phrasal standards. Brucart (2003) also picks up on this point and argues in favour of Sáez´s proposal. As we will see next, I will go a step further and propose that sentences like (15) and (16) in Spanish involve phrasal standards in which, crucially, the complement of the standard marker *que* does not conform to the semantic parameter previously described in the literature. These standards are [-CLAUSAL], and consequently, do not involve degree abstraction. They neither denote degrees directly (they are not measure phrases or degree phrases) nor a referential (*e* type) individual. I will refer to inequality comparatives like (15) and (16) as *more N than N* comparatives.

(15) Más   mujeres [que *hombres*] fueron a   la   manifestación.
       More women [than men   ] went to   the march.
       'More women than_que men attended the march.'

(16) La huelga contra la   LOMCE arrastra a   más   alumnos [que *profesores*].[7]
       the strike against the LOMCE drags   DOM more students [than teachers   ]
       'The strike against the LOMCE drags more students than_que teachers.'

The acceptability of (15) contrasts with (17a), where the standard is clearly clausal. Clausal standards in Spanish cannot be linearized in a position between the subject and the verb of the matrix clause (see (17a-b)). The contrast in grammaticality between (15) and (17a) will lead us to defend a phrasal analysis for the standard in (15).

(17) a.   *Ayer,   más mujeres [que hombres (fueron) la   semanapasada] fueron a
              yesterday more women [than men   went   the week   past   ] went to
              la   manifestación.
              the   march.
              Lit. 'Yesterday, more women than men (attended) last week attended the march.'
       b.   Ayer,   más mujeres fueron a   la   manifestación [que hombres (fueron)
              yesterday more women went   to the march   [than men   went
              la   semana   pasada].
              the   week   past   ]
              'Yesterday, more women attended the march than men (did) last week.'

Furthermore, note that not just any phrasal element can act as a standard and be linearized between the subject and the verb of a clause, as (18a-b) show. In these sentences, the complement of the standard marker is a temporal adverb, possibly modifying a verb that has been elided due to identity with the verb in the matrix clause. Thus, (18b) might be a case of a reduced clausal standard which looks phrasal on the surface due to elision of some material within the clause.

(18) a.   *Hoy más mujeres [que ayer]   han ido a   la   manifestación.
              todaymore women   [than yesterday] have gone to   the march
              Lit: 'Today more women than_que yesterday have attended the march.'
       b.   Hoy   más mujeres han ido a   la   manifestación [que ayer].
              today more women have gone to   the march   [than yesterday]
              'Today more women have attended the march than_que yesterday.'

What we learn from the contrast between (17) and the ungrammatical (18a) is that the ban on clausal standards being linearized between the subject and the verb of the matrix clause is not due to a constraint that only allows simple, phrasal standards in that position, but to a ban on any type of clausal standards (reduced and unreduced) in subject positions.
    The schematic bracketing proposed for (12=19a), is presented in (19b):

---

[7] Headline from an article in El Diario Montañés (09-03-2017)
<http://www.eldiariomontanes.es/cantabria/201703/08/convocantes-huelga-educativa-llaman-20170308211154.html>

(19) a.   Más   mujeres que hombres fueron a   la manifestación.   [- CLAUSAL]
          More   women   than men   went to the march.
          'More women than_que men attended the march.'
     b.   [_DP Más mujeres [_thanP que hombres]]]

Furthermore, *more N than N* comparatives such as (19) in Spanish present two other interesting properties, which have gone unnoticed in the literature: (i) they are not restricted to subject positions only; and (ii) they are subject to an adjacency restriction between the element modified by the comparative marker *más* and the standard of comparison. These two properties are exemplified in (13), repeated below as (20a). In (20a), the comparative marker *más* is modifying the bare plural *alumnos*, the object of the sentence, which is introduced by the preposition *a*.[8] In this case, the standard of comparison is the NP *profesores*. If we extrapose the standard and introduce some adjuncts between the modified nominal and the standard as in (20b), the use of the DOM marking preposition becomes obligatory. This adjacency restriction supports the bracketing proposal in (20c) and the general schematic analysis proposed for these structures in (21).

(20) a.   La   huelga contra la LOMCE arrastra[a   [más   alumnos [que profesores]]].
          the   strike against the LOMCE drags   [DOM[more   students [than teachers]]]
          'The strike against the LOMCE drags more students than_que teachers.'
     b.   La   huelga contra la LOMCE arrastra[a   más   alumnos]a la calles
          the   strike against the LOMCE drags   [DOM more students]to the streets
          cada semana[que *(a)   profesores].
          every week   [than *(DOM) teachers]
          'The strike against the LOMCE drags more students to the streets every week than_que (it does) teachers.'
     c.   [_PP a [_DP más [_NP alumnos] [_thanP que [_NP profesores]]]]
     d.
(21) [_DP more [_NP Xs] [_thanP than [_NP Ys]]

In sum, the properties of standards of comparison in comparatives with a structure like the one in (21) − that of sentences such as (15) and (16) − cannot be accounted for by the semantic [INDIVIDUAL]/[DEGREE] parameter, because the standards in these cases do not involve degree abstraction, nor do they denote either degrees directly (they are not a MPs), or a referential individual (they are not DPs).

2.2. Event quantification comparatives

We find another case that cannot be accounted for by previous semantic analyses in one kind of clausal comparatives in Spanish, which I will refer to as *event quantification* comparatives. First, consider the pair of comparative sentences which clearly present clausal standards with contrasting word orders in (22).

(22) a.   Jon compró más libros que *Ana leyó cómics.*

---

[8] This is so because Spanish is a Differential Object Marking (DOM) language.

Jon bought   more books than Ana read comics
'Jon bought more books than_{que} Ana read comics.'
b.   Jon compró más libros que *cómics leyó Ana.*
Jon bought   more books than comics   read Ana
(Lit.) 'Jon bought more books than_{que.inv} comics read Ana.

In these cases, it is clear that the standard marker *que* takes a fully clausal element in its complement position. I will argue that this word order difference is not trivial since, in fact, sentences (22a-b) present radically different syntactic and semantic properties (see *Table 1* for a summary of the differences).

A.   Sentence 0 is sensitive to coordination constraints, while 0 is not. The fact that sentences like (22a) are subject to coordination constraints was already attested by Sáez (1992) and some of his arguments will be reviewed in the following paragraphs.
B.   Both sentences have different interpretations. The (b) version with OVS word order compares the relative quantities of different sorts of objects. In (22b) in particular, the number of books is compared to the number of comics. But crucially, in the (a) version, I will argue that the relative frequencies of two types of events - the number of instances in which those events have occurred - are compared. In this case, *more* seems to quantify over number of *events* and it is thus comparing frequencies or pluralities of events. Hence, sentence (22a) has an interpretation similar to 'There have been more instances of Jon buying books than instances of Ana reading comics'.

Comparative sentences such as (22b) are known as *subcomparatives.* According to Bresnan (1973) and many other authors after her (cf. Sáez 1992, Brucart 2003, Reglero 2007, for Spanish *subcomparatives*), the underlying representation of (22b) contains a variable ranging over amounts or degrees of cardinality. Moreover, one of the most striking properties of Spanish *subcomparatives* is their unusual surface word order, with the object (cf. (22b)) or a gradable predicate (23) first and subject-verb inversion. This inversion pattern is also present in *wh*-movement contexts, such as questions (Torrego 1984).

(23) Esta mesa es más larga que *ancha es esa   puerta.*
This table is   more long than wide   is   that door
'This table is longer than that door is wide.'

In contrast, comparatives with a clausal standard that presents the canonical word order (SVO) will be argued to conform another class of comparatives with a different underlying syntactic structure, as Sáez (1992) already suggested, and , interestingly, also to a different interpretation. There are two novel observations that, as far as I know, have not received a formal explanation in the literature on Spanish comparatives. The first is the interpretive contrast between (22a) and (22b); and the second is the fact that comparatives with a clausal standard that has the canonical word order (SVO) as in (22a) involve a comparison on the relative frequencies of two types of events, namely, the number of instances in which those events have occurred.

The interpretive difference that I am alluding to distinguishes (i) comparison of amounts of objects (degrees of cardinality) or degrees of a gradable predicate in *subcomparatives* like (22b) and (23), from (ii) comparison of number of instantiations of some events (degrees of

frequency) in *event quantification* comparatives (22a). This distinction is a very subtle one. In order to check that the above intuitions are correct, we can test some predictions derived from these hypotheses.

In (24) and (25) below, two scenarios are created in which one of the interpretations (either the degree or the event quantification reading) is very salient and the second reading is quite unnatural. First, in (24), a context in which the *event quantification* interpretation is more salient is created. Here, the comparison is located at some specified interval of time, note that the verbs in the matrix clause and the standard clause are stage-level predicates (the verb *estar*) and closed-scale, absolute adjectives are employed (cf. Kennedy & McNally 2005, Gumiel & Pérez-Jiménez 2012). If in clausal comparatives with an SVX word order in the standard *more* is quantifying over number of events, then, our prediction is that in a clausal comparative with the properties of (24), the SVX word order will be preferred over an XVS one. As the acceptability contrast between (24a) and (24b) shows, this first prediction is borne out: the *event quantification* interpretation is associated to the use of an SVX word order in the standard of comparison.

(24) a.   Este mes,   Jon ha estado más enfermo que *Mikel ha estado sano.*[9]
           this month Jon has been   more sick   than Mikel has been   healthy
           'This month, Jon has been sick more (times) than Mikel has been healthy.'
      b.   *Este mes,   Jon ha estado más enfermo que *sano   ha estado Mikel.*
            this month Jon has been   more sick   than healthy has been   Mikel
           'Jon has been more sick than Mikel has been healthy.'

Second, in (25), the degree comparison interpretation is made more salient by using an individual-level, gradable predicate (more specifically, a open-scale, relative adjective combined with *ser*) in the standard. If the interpretation of clausal comparatives with an XVS word order in the standard is that of a comparison between two degrees associated to two different gradable adjectives, as proposed, I also predict that, in a context such as (25), an XVS word order in the standard should be preferred over the canonical SVX word order. This second prediction is also borne out, as shown by the contrast in grammaticality between (25a), with an SVX word order, and (25b), with an XVS word order.

(25) a.   *Jon es más comunista que la sangre es roja.[10]
            Jon is more communist than the blood is   red
           'Jon is more communist than the blood is red.'
      b.   Jon es más comunista que roja es la sangre.
           Jon is more communist than red   is   the blood
           'Jon is more communist than blood is red.'

In fact, the predicate-subject-verb word order is also preferred in typical cases of subcomparatives, as in (23), where two degrees are compared.

---

[9] As a reviewer suggests, (24a) might also contribute a duration interpretation ("...has been sick for a longer period of time than Mikel has been healthy.') in addition to the event quantification interpretation. Since we are focused on understanding the constrast between (22a) and (22b), and in (22a) no such interpretation is available, I will not delve into this second reading that could be associated with (24a).

[10] Thanks to Myriam Uribe-Etxebarria for offering this example.

Summarizing, what we have seen so far is that clausal comparatives with an XVS word order compare relative quantities of different sorts of objects or degrees associated to different gradable predicates. These comparatives have been referred to as *subcomparatives.* In contrast, clausal comparatives with an SVX word order compare the number of instances in which two types of events have occurred. I have proposed the label *event quantification* comparatives for these latter cases.

With respect to their syntactic differences, in his 1992 paper on Spanish comparatives, Sáez already argued extensively that comparatives with an SVO clausal standard are subject to coordination constraints, whereas comparatives with an OVS standard are not. A justification of the coordination analysis immediately follows.

For limitations of space, I will only review three of Sáez´s arguments that support the claims that these two types of comparatives have different syntactic structures and that comparatives like that in (19a) should be analysed as involving coordination.[11]

First, what I am calling *event quantification* comparatives, i.e. sentences like (22a) and (24a), are subject to a constraint on extraction from the standard clause (see (26b)), but allow Across-the-Board (ATB) extraction from both clauses (see (26a)), according to Sáez (1992). On the other hand, as the author argues, *subcomparatives* such as (22b) and (23) do not present this constraint, and extraction from the standard clause is available (cf. (27)). The same restrictions that apply to *event quantification* comparatives also apply to general cases of coordination, as the sentences in (28) show.

Event quantification:
(26) a.   ¿A quién compró Pedro más manzanas _ que   vendió Juan peras _?
             to   whom bought Pedro more apples   than_que sold   Juan pears ?
             Lit. 'To whom did Pedro buy more apples _ than John sold pears _?'
       b.   *¿ A quién compró Pedro más manzanas _ que vendió Juan peras a Luis?
             to whom bought Pedro more apples _   than sold   Juan pears to Luis?
             'To whom did Pedro buy more apples _than_que sold Juan pears to Luis?'

Subcomparative:
(27) ¿A quién   compró Pedro más manzanas _ que peras vendió Juan a Luis?
        to whom   bought Pedro more apples   _ than pears sold   Juan to Luis
        Lit.'To whom did Pedro buy more apples _ than_que pears sold Juan to Luis?'

Coordination:
(28) a.   ¿ A quién   compró Pedro manzanas _ y   vendió Juan peras _?
             to whom   bought Pedro apples   _ and   sold   Juan pears _
             Lit. 'To whom did Pedro buy apples _ and sold Juan pears _?'
       b.   *¿ A quién compró Pedro manzanas _ y   vendió Juan peras a   Luis?
             to whom bought Pedro apples   _ and sold   Juan pears   to Luis
             'To whom did Pedro buy apples _ and sold Juan pears to Luis?'

Thus, in comparatives involving event quantification the standard marker *que* seems to behave as a conjunction (*que_&*) coordinating the matrix and the standard clauses, while it behaves as a preposition (*que_prep*) introducing an embedded clause in subcomparatives (Sáez 1992).

---

[11] The reader is encouraged to consult Sáez (1992) for further evidence supporting this point.

Second, another interesting property that distinguishes *event quantification* comparatives from *subcomparatives* in Spanish is the availability of long distance movement (Sáez 1992). *Subcomparatives* allow comparison with a standard that is deeply embedded, as is the case of 'pears' in sentence (29). Nevertheless, sentence (30) does not have the interpretation that there were more instances of Juan buying apples than the number of instances of Luis selling pears that María thought there were.

Subcomparative:
(29) Juan compró más manzanas que peras se piensa María que vendió Luis.
      Juan bought more apples   than pears thinks María that sold Luis
      Lit. 'Juan bought more apples than_que pears thinks María that sold Luis.'

Event quantification:
(30) *Juan compró más manzanas que María se piensa que Luis vendió peras.
       Juan bought more apples   than María thinks   that Luis sold   pears
      'Juan bought more apples than_que María thinks that Luis sold pears'.

A third piece of evidence that distinguishes these two types of comparatives in Spanish and which argues in favour of the coordination analysis of *event quantification* comparatives is the constraint on subjunctive mood shown in (31) (cf. Sáez 1992). This constraint is also present in coordinated sentences in Spanish (see (32)), but not in *subcomparative* sentences like (33), nor in comparatives in which the standard of comparison is introduced by the preposition *de*, as in the example in (34). The following examples have been slightly modified from those in Sáez's paper in order to have the most similar possible minimal pairs.

Event quantification:
(31) *Juan   ha comprado más manzanas que Ana haya   podido vender peras.
       Juan   has bought   more apples   than Ana has_subj could sell   pears
      'Juan has bought more apples than_& Ana could_subj sell pears.'

Coordination:
(32) *Juan ha comprado manzanas y   Ana haya   podido vender peras.
       Juan has bought   apples   and Ana has_subj could sell   pears
      'Juan has bought apples and Ana could_subj sell pears.'

Subcomparative:
(33) Juan ha comprado más manzanas que peras haya   podido vender Ana.
      Juan has bought   more apples   than pears has_subj could sell   Ana.
      'Juan has bought more apples than pears could_subj sell Ana.'

Comparative with de_prep *standard*:
(34) Juan compró más manzanas de   las que Ana haya podido vender.
      Juan bought more apples   than the that Ana has_subj could sell
      'Juan bought more apples than_de the ones that Ana could_subj sell.'

Hence, as Sáez (1992) argues, *event quantification* comparatives seem to pattern with cases of clausal coordination, and *subcomparatives* pattern with comparatives with the prepositional standard *de* and introduce an embedded clause.

The syntactic and semantic differences between the two types of comparatives with a clausal standard in Spanish are summarized in *Table 1*.

	Subcomparatives: *que*$_{prep}$ **XVS**	**Event quantification comparatives:** *que*$_{\&}$ **SVX**
Syntax	As argued by Sáez (1992), *subcomparatives* do not display coordination constraint effects: ✓ any kind of extraction ✓ long distance movement ✓ subjunctive mood In the standard clause, there is *wh*-movement of the degree to the left periphery of the clause and pied-piping of the restrictor (cf. Brucart 2003). The degree abstraction is triggering the subject-verb inversion within the embedded clause.	As argued by Sáez (1992), the matrix and the standard clause are coordinated. These comparatives are subject to coordination constraints: ✗ only ATB extraction ✗ long distance movement ✗ subjunctive mood New proposal: No degree abstraction is involved in the standard clause (evidenced by the lack of subject-verb inversion in the embedded clause).
Semantics	In *subcomparatives*, the standard denotes a degree directly and thus degrees of some gradable predicate or degrees of cardinality (number of objects) are compared (Bresnan 1972).	New proposal: *event quantification* comparatives involve comparison of frequencies or "times" of an event happening, that is, they compare the number of sub-events.

Table 1: Two types of comparatives with clausal standards in Spanish

Finally, if we assume that what triggers subject-verb inversion in Spanish *subcomparatives* is *wh*-movement of a degree element to [Spec, CP], then the lack of subject-verb inversion in *event quantification* comparatives should be interpreted as reflecting the fact that they do not involve degree abstraction. This hypothesis is supported by the observation that long distance *event quantification* comparatives are not possible (cf. (30)).

Hence, *event quantification* comparatives are another case of standards of comparison that do not conform to the [INDIVIDUAL]/[DEGREE] semantic parameter proposed in the literature. As was the case for *more N than N comparatives* presented in Section 2.2., *event quantification* comparatives do not involve a directly degree-denoting standard (there is no degree abstraction, as evidenced by the lack of subject-verb inversion in the standard clause and the lack of a long distance event interpretation) and they are not individual-denoting

standards either. This means that we need to modify the semantic parameter in order to account for the variation observed in these standards of comparison.

Following the logic of the syntactic parameter presented in Section 1.1, I would like to propose a *semantic parameter* that distinguishes directly degree-denoting standards -- that is, those standards of comparison that present a MP or involve some kind of degree abstraction (*wh*-movement or a degree relative clause) and are thus [+ DEGREE] denoting - from standards that do not denote degrees directly and which will be represented as [- DEGREE] denoting.

The new classification of Spanish comparatives I propose is given in *Table 2* and is based on two independent parameters: a syntactic parameter [-/+ CLAUSAL] and a revised semantic parameter [-/+ DEGREE].

<table>
<tr><td rowspan="2" colspan="2"></td><th colspan="2">SYNTACTIC</th></tr>
<tr><th>- CLAUSAL</th><th>+ CLAUSAL</th></tr>
<tr><th rowspan="3">SEMANTIC</th><td>- DEGREE</td><td>Spanish (*que* phrasal)</td><td>Spanish (*que* event quantif.)</td></tr>
<tr><td rowspan="2">+ DEGREE</td><td>QP or amount RC</td><td>*wh*-movement</td></tr>
<tr><td>Spanish (*de*)</td><td>Spanish (*que* subcomparative)</td></tr>
</table>

Table 2: Parametric variation on the expression of standards of comparison in Spanish

3. Analysis

Following the work by Bartsch & Vennemann (1972) or Kennedy (1999), among others, in adjectival comparatives I assume a measure function analyisis of gradable predicates, as in (35).

(35) $[\![\text{tall}_{\mu \, <e,d>}]\!] = \lambda x.\text{tall}(x)$

In nominal comparatives, *more* is analysed as an analytic comparative marker consisting of the quantity word *many* and the comparative marker, following Bresnan (1973), as illustrated in (36).

(36) *more* $= \mu + \text{er}$

Following the counting function analysis of the quantity word *many* proposed for English (cf. Schwarzschild 2006, Kennedy 2002, 2007, Rett forth., Solt 2015, a.o.), I assume that in nominal comparatives Spanish *más* also includes a quantity word that behaves as an hybrid between a quantifier and a gradable predicate, represented as 'μ', as illustrated in (37).

(37) $[\![\mu_{<et, \, d>}]\!] = \lambda X.\mu(X)$

With these premises in mind, a novel formal analysis of Spanish inequality comparatives is presented in Sections 3.1, 3.2 and 3.3 that accounts for the syntactic and semantic properties examined in Section 2.

De standards in Spanish take either a MP or a degree relative clause (cf. Sáez 1992, Grosu & Landman 1998, Gutiérrez-Rexach 1999). Following the proposed parameters, these standards are [+ DEGREE] denoting in the semantics and have a [- CLAUSAL] structure.

(38)   a.   Ana compró más libros de   *los   que   ves ahí.*
              Ana bought   more books than the   that see there
              'Ana bought more books than$_{de}$ you see there.'
       b.   Ana compró más libros de [$_{QP}$ los [μ ~~libros~~]$_i$ [$_{CP}$ [μ ~~libros~~]$_i$ que *pro$_{you}$* ves ahí [μ ~~libros~~]$_i$]].[12,13]
              'Ana bought more books than [$_{QP}$ the [μ ~~books~~]$_i$ [$_{CP}$ [μ ~~books~~]$_i$ that *pro$_{you}$* see there [μ ~~books~~]$_i$]].'

Following McConnell-Ginet's (1973) idea that all inequality comparatives involve the presence of a gap - that is, the difference between the two compared values, which equals a set of degrees or an interval (cf. Schwarzschild & Wilkinson 2002, Schwarzschild 2005) - *más libros de los que ves ahí* in (38) would have the denotation in (39).[14]

(39)   ⟦más N de *d* ⟧ = $\lambda\mu$ λP λd $\exists$X $\exists$I [P(X) & I = μ(X) - d & I $\subseteq$ (0,+∞)]

The observation that *de* comparatives in Spanish are closely related to magnitudes has always been present in traditional grammars of Spanish. The representation in (39) of *más de* comparatives supposes a novel approach to the formalization of that long-standing observation. See also Mendia (forth.) for independent evidence in favor of the degree-denoting analysis of *más de* comparatives.

The data examined in Section 2 appeals to a distinction between two main types of *que* standards. In particular:

---

[12] Sáez (1992), Grosu & Landman (1998) and Gutiérrez-Rexach (1999), a.o., offer different analyses for degree relatives. The comparison of these three proposals goes beyond the scope of this paper. The reader is referred to these works for related discussion.

[13] We can find examples were 'μ' seems to be explicit in sentences like (i) from the newspaper *Diario de León* (05/06/2013).

    (i)   Esto es un crimen más *de   los muchos* que hay   en este país.
        this is a crime more than the many   that happen in this country
        'This is one more crime of (than$_{de}$) the many that happen in this country.'

[14] The analysis of comparatives presented in this paper makes use of non-empty intervals to describe the difference between the compared elements. Nevertheless, nothing hinges on this particular approach to inequality comparatives, and a more traditional analysis in terms of a greater than (>) relation could also be applied in (39). However, the possibility of licensing a differential is one of the main features that distinguishes inequality comparatives from other comparatives and degree constructions. The application of an analysis in terms of intervals of difference highlights the importance of this crucial component of inequality comparatives.

A.   Conjunctive *que*, which introduces [- DEGREE] denoting standards that can be either [-CLAUSAL] or [+CLAUSAL].
B.   Prepositional *que,* which introduces [+ DEGREE] denoting and [+CLAUSAL] standards.

Those standards that are non-directly degree-denoting [- DEGREE] involve a coordinated structure (Sáez 1992). In these cases, I propose that a measuring function (either the gradable predicate or a μ counting function) is applied to both coordinates to obtain the degree argument that sets the standard of comparison. In the English phrasal version of example (15), repeated below as (40), a function from plural objects to degrees of cardinality (μ) is applied to both coordinates, i.e. *women* and *men*. A schematic structure is proposed in (41), and (42) to illustrate its semantic representation.

More N than N: nominal coordination in a comparative structure
(40)   More women than men attended the march.

(41)   [-er [μ women] than [μ men]] attended the march.

(42) $\exists X \exists Y \exists I$ [women(X) & men(Y) & I = $\mu(X)$ - $\mu(Y)$ & I $\subseteq$ (0,+∞)]
       There are some X, some Y and an interval I, such that X are women and Y are men, and the number of men minus the number of women equals a non-empty interval (a set of degrees).

In clausal standards, following a neo-Davidsonian quantificational representation of events, the sentence in (22a), repeated below in (43), would be assigned the representation in (44). I leave the specifics on the syntactic structure of *event quantification* comparatives such as (43) for future research.

Event quantification comparatives:
(43)   Jon compró más libros que Ana leyó cómics.[15]
       Jon bought more books than Ana read comics
       'Jon bought more books than$_{que}$ Ana read comics.'

(44)   There is an event e_1, whose sub-events consist of book-buyings by Jon, and another event e_2, whose sub-events are comic-readings by Ana, and the number of e_1 sub-events minus the number of e_2 sub-events equals a non-empty interval.

3.3. Spanish: que *subcomparatives*

Clausal standards introduced by *que* in which the subject and the verb of the embedded clause are inverted involve a degree abstraction operation with *wh*-movement (Sáez 1992, Brucart 2003). Consequently, I classify *subcomparatives* as being [+ DEGREE] [+ CLAUSAL].

---

[15] As a reviewer kindly indicates, frequency comparatives in English as in sentence (i) appear similar to the Spanish [- DEGREE,+ CLAUSAL] event quantification comparatives:
       (i)   John buys books more than Mary reads comics.
I leave the examination and comparison of English examples like (i) with Spanish comparatives like (22a) for further research.

In particular, I defend that the standard of comparison in Spanish *subcomparatives* presented in (22b/45) behaves similarly to that of a free relative clause (cf. Sáez 1992) in the sense that it denotes a maximal degree. Consequently, despite their syntactic differences, *subcomparatives* and comparatives where the standard is introduced by the preposition *de* are similar in that, in both cases, *más* takes a degree-denoting standard of comparison, as proposed in (39) and illustrated in (46).

(45)   a.   Jon compró más libros que cómics leyó Ana.
              Jon bought more books than comics read Ana
              'Jon bought more books than Ana read comics.'
        b.   Jon compró más libros que [$_{CP}$ [μ cómics]$_i$ leyó Ana t$_i$].
              'Jon bought more books than [$_{CP}$ [μ comics]$_i$ read Ana t$_i$].'

(46) $\llbracket$más N que d $_{subcomparative}$$\rrbracket$ = $\lambda\mu$ λP λd $\exists$X $\exists$I [P(X) & I = μ(X) - d & I $\subseteq$ (0,+∞)]

3.4. Cross-linguistic perspective

The main questions that this paper addresses are the following: how can we derive a compositional analysis that accounts for the syntactic and semantic properties of standards of comparison? And what are the constraints that apply to standards of comparison cross-linguistically?

The recent literature on standards of comparison establishes the locus of cross-linguistic variation in a syntactic *phrasal* vs. *clausal structure* distinction, and a semantic *individual-denoting* vs. *degree-denoting* distinction.

Looking at data from Spanish comparatives, I have shown that the proposed semantic *individual/degree* parameter of variation is insufficient to account for the observed data in Spanish, and a refinement on this parameter has been proposed.

Specifically, I have shown several examples of standards of comparison which are not (i) degree-denoting (that is, the standard of comparison does not include a phrasal structure with a MP or a degree relative clause, nor a clausal structure with degree abstraction); nor (ii) individual-denoting. For this reason, the semantic parameter has been redefined as a [-/+ DEGREE] parameter, that distinguishes directly degree-denoting standards [+ DEGREE] from standards that do not denote degrees directly [- DEGREE], which can denote individuals, properties or events, for example.

Furthermore, I have proposed an analysis where, in order to obtain the degree that establishes the referent for the comparison in [- DEGREE] denoting standards, a measure function (a gradable predicate or a counting function *many* or *much*) is applied to the element in the standard of comparison.

In sum, in order to answer the main questions discussed in this paper, I have proposed that standards of comparison cross-linguistically can be classified with respect to two independent parameters: a syntactic one [-/+ CLAUSAL] and a semantic one [-/+ DEGREE]. *Table 3* illustrates the parametric variation on the expression of standards of comparison with some sample languages. The cross-linguistic data for this classification has been collected from Hofstetter (2009), for Turkish; Sudo (2015), for Japanese; Heim (1985) and Kennedy (1999), for English; and Donati (1997) and Belleti (1991-1995), for Italian. The classification of Spanish comparatives has been implemented according the analysis presented in Section 3.

The classification of Basque comparatives is a tentative approach to describe standards of comparison in this language. Nevertheless, a deeper analysis of comparatives in this language is still necessary.

		SYNTACTIC	
		- CLAUSAL	**+ CLAUSAL**
SEMANTIC	**- DEGREE**	Turkish (ABL), Japanese (yori), Basque (baino), English (than), Italian (che), Spanish (que)	Spanish (que), Italian (che)
	+ DEGREE	QP or amount RC	*wh*-movement
		Turkish (ABL), Japanese (yori), Basque (baino), English (than), Italian (di), Spanish (de)	English (than), Spanish (que$_{inversion}$), Italian (di$_{inversion}$)

Table 3: Parametric classification of standards of comparison

This novel parametric classification should improve our understanding of cross-linguistic variation of inequality comparatives, as well as setting up the basis for a compositional analysis of comparative structures.

4. Conclusion

In order to elaborate accounts of comparative structures as a whole, a deeper understanding of the individual pieces that compose the puzzle that comparatives represent is necessary. With this goal in mind, I have focused on one of the main components of these structures: the standard of comparison.

Two independent parameters have been argued to be necessary to account for the properties of standards of comparison cross-linguistically: a syntactic one [-/+ CLAUSAL], and a semantic one [-/+ DEGREE] denoting. In particular, Spanish standards of comparison can be classified into three types:

A) Phrasal *de* standards that take either a MP or a degree relative clause as their complement and denote degrees [+ DEGREE] [- CLAUSAL];
B) *Que* standards that involve coordinated structures and are not directly degree-denoting [- DEGREE]. These can be either [- CLAUSAL] or [+ CLAUSAL];

C)   Clausal *que* standards with subject-verb inversion in the standard clause that involve (*wh*) operator movement and denote degrees [+ DEGREE] [+ CLAUSAL].

Finally, the Spanish data also supports those analyses which argue that the comparative meaning depends both on the comparative marker (*más*) and the standard marker (*que* or *de*).

Acknowledgements

Laura Vela-Plo
University of the Basque Country UPV/EHU
laura.vela@ehu.eus
https://metaaprendizaje.academia.edu/LauraVelaPlo

References

Bacskai-Atkari, J. (2014). Cyclical change in Hungarian comparatives. *Diachronica* 31:4, pp. 465-505.
Bale, A. C. (2008). A universal scale of comparison. *Linguistics and Philosophy* 31:1, pp. 1-55.
Bartsch, R. & T.Vennemann. (1972). Relative adjectives and comparison. *UCLA papers in syntax* 2, pp. 80-107.
Beck, S., T. Oda & K. Sugisaki. (2004). Parametric variation in the semantics of comparison: Japanese vs. English. *Journal of East Asian Linguistics* 13, pp. 289–344.
Belletti, A. (1991-1995). Le frasi comparative. In Renzi, L., G. Salvi & A. Cardinaletti (eds.), *Grande grammatica italiana di consultazione*, Il Mulino, Bologna.
Bhatt, R. & S. Takahashi. (2007). Direct Comparisons: Resurrecting the Direct Analysis of Phrasal Comparatives. [Paper presented at *SALT 17*]. University of Connecticut, May. people.umass.edu/bhatt/papers/salt-comparatives-talk.pdf
Bhatt, R. & S. Takahashi. (2008). When to Reduce and When not to: Crosslinguistic Variation in Phrasal Comparatives. *Glow XXXI*, Newcastle Univesity, UK.
Bhatt, R., & S. Takahashi. (2011). Reduced and unreduced phrasal comparatives. *Natural Language & Linguistic Theory* 29:3, pp. 581-620.
Bresnan, J. W. (1972). Theory of complementation in English syntax. [PhD thesis]. Massachusetts Institute of Technology.
Bresnan, J. W. (1973). Syntax of the comparative clause construction in English. *Linguistic inquiry* 4:3, pp. 275-343.
Brucart, J. M. (2003). Adición, sustracción y comparación: un análisis composicional de las construcciones aditivo-sustractivas del español. *Actas del XXIII CILFR* 1, pp. 11-60.
Brucart, J. M. (2009). Patrones formales e interpretación: el funcionamiento de más en español. *Estudios de linguistic* 1, pp. 13-44.
Chomsky, N. (1977). On Wh-Movement. In Culicover, P. W., T. Wasow & A. Akmajian (eds.), *Formal Syntax*, Academic Press, New York, pp. 71–132.
Donati, C. (1997). Comparative clauses as free relatives. *Probus* 9, pp. 145-166.
Eguren, L., Soriano, O. F., & Mendikoetxea, A. (eds.). (2016). *Rethinking Parameters*. Oxford University Press.
Fu, Y. C. (1978). Comparative structures in English and Mandarin Chinese. [PhD thesis]. University of Michigan.

Grosu, A. & F. Landman. (1998). Strange relatives of the third kind. *Natural Language Semantics* 6:2, pp. 125-170.

Gumiel, S. & I. Pérez-Jiménez, I. (2012). Aspectual composition in" ser/estar+ adjective" structures: adjectival scalarity and verbal aspect in copular constructions. *Borealis–An International Journal of Hispanic Linguistics* 1:1, pp. 33-62.

Gutiérrez Ordóñez, S. (1994). *Estructuras comparativas*. Arco Libros, Madrid.

Gutiérrez-Rexach, J. (1999). The structure and interpretation of Spanish degree neuter constructions. *Lingua* 109:1, pp. 35-63.

Hankamer, J. (1973). Why there are two *than*'s in English. In C. Corum, T.C. Smith- Stark & A. Weiser (eds.), *Papers from the 9th regional meeting of the Chicago Linguistics Society,* Chicago Linguistics Society, Chicago, Illinois, pp. 179-191.

Heim, I. (1985). Notes on comparatives and related matters. [Ms]. University of Texas, Austin.

Hoeksema, J. (1984). To be continued: the story of the comparative. *Journal of Semantics* 3:1-2, pp. 93-107.

Hofstetter, S. (2009). Comparison in Turkish: a rediscovery of the phrasal comparative. In *Proceedings of Sinn und Bedeutung* 13, pp. 187-202.

Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. Garland, New York.

Kennedy, C. (2002). Comparative deletion and optimality in syntax. *Natural Language & Linguistic Theory* 20:3, pp. 553-621.

Kennedy, C. (2006). Semantics of comparatives. In Allen, K. (ed.), *Lexical and Logical Semantics section of the Encyclopedia of Language and Linguistics*, Elsevier, Oxford.

Kennedy, C. (2007). Modes of comparison. In *PCLS 43* 1:1, pp. 141-165.

Kennedy, C. & L. McNally. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, pp. 345-381.

Martínez, J. A. (1987). Construcciones y sintagmas comparativos en el español actual. *In Memoriam Inmaculada Corrales*, pp. 319-336.

McConnell-Ginet, S. (1973). Comparison constructions in English. [PhD thesis]. University of Rochester.

Mendia, J.A. (forth). One More Comparative. [Ms]. University of Massachusetts, Amherst, September 2016.

Napoli, D. J. (1983). Comparative ellipsis: A phrase structure analysis. *Linguistic Inquiry* 14:4, pp. 675-694.

Pancheva, R. (2006). Phrasal and clausal comparatives in Slavic. In Lavine, E (ed.), *Formal approaches to Slavic linguistics 14*: *The princeton meeting*, Michigan Slavic Publications, Ann Arbor, pp. 236–257.

Reglero, L. (2007). On Spanish comparative subdeletion constructions. *Studia linguistica* 61:2, pp. 130-169.

Rett, J. (forth). The semantics of many, much, few, and little. [Ms]. University of California, Los Angeles, August 2016. http://www.linguistics.ucla.edu/people/Rett/Rett%202017%20quantity%20words.pdf

Romero Cambrón, Á. (1998). *Historia sintáctica de las construcciones comparativas de desigualdad*. Ediciones de la Universidad de Castilla-La Mancha, Cuenca.

Rullmann, H. (1995). Maximality in the Semantics of Wh-Constructions. [PhD thesis]. University of Massachusetts, Amherst.

Sáez, L. A. (1992). Comparison and Coordination. *Anuario Del Seminario de Filología Vasca"Julio de Urquijo", Syntactic Theory and Basque Syntax*, pp. 387–416.

Schwarzchild, R. & K. Wilkinson. (2002). Quantifiers in comparatives: A semantics of degree based on intervals. *Natural language semantics* 10:1, pp. 1-41.

Schwarzschild, R. (2005). Measure phrases as modifiers of adjectives. *Recherches linguistiques de Vincennes* 34, pp. 207-228.

Schwarzschild, R. (2006). The role of dimensions in the syntax of noun phrases. *Syntax* 9:1, pp. 67-110.

Shimoyama, J. (2012). Reassessing crosslinguistic variation in clausal comparatives. *Natural Language Semantics* 20:1, pp. 83-113.

Solt, S. (2015). Q-adjectives and the semantics of quantity. *Journal of Semantics* 32:2, pp. 221-273.

Sudo, Y. (2015). Hidden nominal structures in Japanese clausal comparatives. *Journal of East Asian Linguistics* 24:1, pp. 1-51.

Torrego, E. (1984). On inversion in Spanish and some of its effects. *Linguistic inquiry* 15:1, pp. 103-129.

von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics* 3, pp. 1-77.

Xiang, M. (2003). A phrasal analysis of Chinese comparatives. *Proceedings from the Annual Meeting of the Chicago Linguistic Society* 39:1, pp. 739-754. Chicago Linguistic Society.

Feature inheritance and the syntax of lexical VV compounds

Ryoichiro Kobayashi

This paper points out a correlation between the presence of productive lexical VV compounds and the lack of object φ-agreement in comparative perspectives. Japanese abounds in lexical VV-compounds (Fukushima 2005), in which neither heads are functional nor auxiliarized unlike syntactic or aspectual compounds (Kageyama 1993). I focus on such VV-compounds, following the assumption in Distributed Morphology that they are formed in Narrow Syntax (Nishiyama 2008). The aim of this study is twofold: Observing seven languages, I demonstrate that a language can form productive VVs if it lacks object-verb φ-agreement. Then, I propose a formal account of the syntax of VV-compounds.

1. Introduction

Japanese abounds in so-called Lexical VV compounds (Kageyama 1993), which are fairly productive unlike in many other languages (Fukushima 2005). It allows many types of combinations as exemplified in (1), which includes even the ones that do not abide by the Transitivity Harmony Principle (Kageyama 1993, 1999, 2016) in (2).

(1) a. tobi-ori   (unergative-unergative)
  jump-drop
  'jump off'

 b. nomi-aruk (transitive-unergative)
  drink-walk
  'drink while walking/go bar-hopping'

 c. naki-haras (unergative-transitive)
  cry-get.swollen
  'cry out'

 d. naguri-koros (transitive-transitive)
  hit-kill
  'beat to death'

 e. ki-kuzure (transitive-unaccusative)
  strike-smash
  'knock to pieces'

 f. tare-nagas (unaccusative-transitive)
  drip-pour
  'drain'

 g. koroge-oti (unaccusative-unaccusative)
  roll-fall
  'roll down'

 *hasiri-korob (unergative-unaccusative)
  run-fall.down
 *korobi-hasir (unaccusative-unergative)
  fall.down-run

(2) The Transitivity Harmony Principle:

    In a VV-compound, if V1 is a verb that has an external argument, V2 must also be a verb that has an external argument, whereas if V1 is a verb that lacks an external argument, V2 must also be a verb that lacks an external argument.

<div align="right">(Nishiyama and Ogawa 2014:88-89)</div>

The definition of Lexical VV-compounds in this article is in (3), which is cited from Kageyama (2016) with a slight modification.

(3) Definition:

    Endocentric VV-compounds that satisfy the non-interruptibility principle of lexical integrity, which behave as indivisible V^0 units in the phrasal syntax.

<div align="right">(Kageyama 2016:278)</div>

The rich presence of Lexical VV compounds in Japanese has sparked numerous studies, both descriptively and theoretically, as well as syntactically and semantically (Kageyama 1993, Fukushima 2005, Yumoto 2005 for Lexical Semantic approaches; Nishiyama 1998, 2008, Nishiyama and Ogawa 2014 among others for Syntactic approaches). However, it has not been answered, nor even ever questioned to the best of my knowledge, why some languages allow such productive VV compounds, while other languages do not. In this paper, I will focus on the cross-linguistic variation of such Lexical VV compounds, and aim to explore a possible account of its presence and absence from comparative perspectives.[1]

    This paper is organized as follows. In section 2, I will analyze data from seven languages to see whether productive VV compounding is possible. I point out a correlation the between presence of Lexical VV compounds and the lack of agreement phenomena. Section 3 proposes a syntactic account of the descriptive generalization in the previous section on the cross-linguistic variation of the Lexical VV-compounds. In section 4, I further touch on the relation between Case system and φ-agreement in the framework of Chomsky (2000, 2008), who argues that Case is licensed as a reflex of φ-feature agreement. Section 5 is an overall summary of the paper.

    Before moving into the body of the paper, some notes on the topics which this paper is not concerned with are in order. There are many aspects one may expect the present paper to deal with, which are empirical and theoretical investigations of VV-compounds, but which in fact it is not concerned with. Three of them that are worth mentioning here are (i) the syntax of Syntactic VV compounds, (ii) the Lexicalist analysis of VV-compounds, and (iii) some exceptional verb-verb sequences that are NOT productive (Lierber 2005:378), to which I now turn.

    VV-compounds are largely divided into two groups: Lexical and Syntactic compounds. It has long been assumed that the former is created in the Lexicon, whereas the latter ones in (4) are derived in the syntactic component (Kageyama 1993, 1999).

---

[1] In this paper, I only focus on nominative-accusative languages and do not discuss the ergative-absolutive or split-ergative languages.

(4) Syntactic VV compounds are usually bi-clausal:

a. aisi-tuduke
   love-continue_{ASP}

b. utai-hazime
   sing-begin_{ASP}

c. kaki-oe
   write-finish_{ASP}

d. tabe-sase
   eat-CAUSE

e. home-rare
   praise-PASS

While some argue that it is still controversial, Nishiyama (1998; 2008) proposes that both the Lexical and Syntactic VV-Compounds are formed in syntax. I basically follow this line of Distributed Morphology-based approach of Nishiyama (1998, 2008), Nishiyama and Ogawa (2014) and many subsequent works, despite their differences in details.

Although the term *Lexical* is confusing, I will continue to use it to refer to the examples listed in (1), which are the main foci of this study. The Syntactic VV-compounds that we are not going to deal with in this paper satisfies the diagnostics in (5).

(5) Syntactic criteria for the lexical-syntactic distinction:
   a. *Soo su*-substitution:
      (i) Lexical: (1a) tobi-ori   *soo si-ori
                         jump-go.down   so   do-go.down

                    (1b) nomi-aruk   *so   si-aruk
                         drink-walk   so   do-walk

      (ii) Syntactic: (4a) aisi-tuduke   soo   si-tuduke
                           love-continue   so   do-continue

                      (4b) utai-hazime   soo   si-hazime
                           sing-begin   so   do-begin

   b. Passivization of V1:
      (i) Lexical: (1c) naki-haras   *nak-are-haras
                         cry-be.swollen   cry-PASS-be.swollen

                    (1d) naguri-koros   *nagur-are-koros
                         punch-kill   punck-PASS-kill

      (ii) Syntactic: (4c) kaki-owar   kak-are-owar
                           write-ASP   write-PASS-ASP

                      (4d) tabe-sase   tabe-rare-sase
                           eat-CAUS   eat-PASS-CAUS

                                                (adapted from Kageyama 2016:278)

Second verbs in Syntactic compounds are analyzed either as a functional or as an auxiliarized element without lexical status (Yashima 2008, Fukuda 2012). Since they show quite different

behaviors from Lexical VV-compounds, and are usually assumed to be derived from biclausal structures (plus restructuring), I do not discuss these Syntactic VV-compounds in this paper.

Let us turn to the second point. Numerous studies on the Lexical VV-compounds are conducted in the framework of the Lexical-Conceptual Structure (LCS: Yumoto 2005). However, since the focus of this paper is on comparative syntax of such compounds, I do not discuss the semantic analysis or those based on LCS here, though they have insightful perspectives (see Nishiyama 2008 for an overview). Furthermore, this paper distinguishes productive Lexical VV-compounds and non-productive exceptional frozen expressions. One may easily think of with some verbal sequences in English such as *dry-clean*, *blow-dry*, *slam-dunk*, *stir-fry*, et cetera. Lieber (1992:80) has shown that compounds of NN, NA, AA and AN combinations are productive in English. However, those containing V as one or both members are frozen expressions (Lieber 2005:378) based on the data in the Corpus of Contemporary American English (COCA) and other corpora (Bauer et al. 2013:451-453, Rochelle Lieber p.c.). In this paper, I focus on whether a language has some syntactic process for productive VV-compounding. We are now ready to procced to the cross-linguistic observations.

2. Endocentric lexical VV compounds beyond Japanese

In this section, we observe six different languages; Korean, Mongolian, Malayalam, Turkish, Bangla, and Ìgbo, to identify whether each one allows Lexical VV-compounds that satisfy the definition in (3), which is repeated here as (6).

(6) Endocentric VV-compounds that satisfy the non-interruptibility principle of lexical integrity, which behave as indivisible V^0 units in the phrasal syntax.

<div align="right">(Kageyama 2016:278)</div>

2.1 Languages with productive lexical VV compounds

Korean, like Japanese, allows productive VV compounds, as in (7). In the same vein, Mongolian (Altaic) and Malayalam (Dravidian) also allow Lexical VV compounds, as illustrated in (8) and (9) respectively. These languages are similar to Japanese in many other respects too, though I do not review them in detail here (see Aoyagi 2011, and series of his works, and Washio 1995 *inter alia*, for further details). [2]

---

[2] *-e* and *-a* are phonological linkers that are inserted in PF (Ko and Sohn 2015). In Korean, some VV-compounds (serialized verbs) are separated by *-se* 'and then' or other particles, as in (i). However, there is another type of compounds that are inseparable in (ii) (Ko and Sohn 2015:83). Although it is controversial whether Korean has inseparable VV-compounds, I assume that it does, given the observations in (ii). See Aoyagi (2014) for details.

(i) a. John-i kaymi-lul palp-a-***se***   cwuk-i-ess-ta.
       J.-NOM ant-ACC trample-LK-SE   die-CAUS-PAST-DECL
       'John trampled an ant to death.'
    b. John-i kaymi-lul palp-a   ***kuphi***   cwuk-i-ess-ta.
       J.-NOM ant-ACC trample-LK   quickly   die-CAUS-PAST-DECL
(ii) a. *John-i Mary-lul kkwulh-e-***se*** anc-hi-ess-ta.
       J.-NOM M.-ACC   kneel-LK-SE   sit-CAUS-PAST-DECL
       'John made Mary kneel down.'
    b. *John-i Mary-lul kwulh-e ***kuphi***   anc-hi-ess-ta.
       J.-NOM M.-ACC   kneel-LK quickly   sit-CAUS-PAST-DECL

(7) a. ttwi-e-nem   b. kwulm-e-cwuk
    jump-go.over   starve-die
    'jump over'   'starve to death'

                                                           (Ko and Sohn 2015)

(8) a. dza:j-ögöx   b. avc-irex
    teach-give   take-bring
    'show'   'bring'

                                                            (Khurelbat 1992)

(9) a. pookuwaan-anuwadicc   b. ār-āy
    go-permit   become.full-search
    'permit leave'   'investigate'

                                                            (Krishnamurti 2003)

2.2 A descriptive generalization

What do these languages with VV-compounds have in common? A striking feature is that these languages all lack grammatical φ-agreement. Mongolian and Malayalam lack φ-feature agreement altogether, just like Japanese and Korean (Fukui 1986 among others), as in (10) and (11).

(10) Mongolian:
    a. Bat-Ø   nama-ig/chama-ig/ter-ig/bid-nig/tanar-ig/ted-nig   har-san.
        B.-NOM   me/you/him/us/you/them-ACC   see-PERF
        'Bat saw me/you/him/us/you/them.'

    b. Bi/Chi/Ter/Bid/Tanar/Ted   Bat-ig   har-san.
        I/You/He/We/You/They   B.-ACC   see-PERF
        'I/You/He/We/You/They saw Bat.'

                                                            (Sakamoto 2011:33)

(11) Malayalam:
    a. Siita   raaman-e   sneehiḱḱunnnu
        Sita   Raman-ACC   love
        'Sita loves Raman.'

    b. raaman   awan-te   bhaarya-ye   sneehiḱḱunnnu
        Raman   he-GEN   wife-ACC   love
        'Raman loves his wife.'

                                                         (Jayaseelan 1999:30-44)

Given these observation, I claim that the absence of φ-features on T or *v* is related to the availability of the Lexical VV-compounding in (12). We are still agnostic at this juncture about whether subject-agreement, object-agreement, or both subject and object-agreement blocks the relevant VV-compounding process in a language. In other words, whether it is φ-features on the functional head, T, or on *v*, or on both heads that are at stake.

(12) Generalization (tentative):
    If a language lacks φ-feature agreement, then it can have Lexical VV-compounds.

2.3 Languages with only subject φ-agreement

Typologically speaking, languages with only object φ-feature agreement are very rare (Moravcsik 1974). In order to further refine the generalization in (12), we will observe languages only with subject φ-agreement: Turkish (Turkic) in (13), Bangla (Indo-Aryan) in (14) and Ìgbo (Niger-Congo) in (15).

(13) Turkish:
    a. Ben   bu   makale-yi   yavaş-yavaş   oku-yacağ-ım
        1.SG   this   article-ACC   slowly-slowly   read-FUT-1.SG
        'I will read the article slowly.'

    b. Biz   her   hafta   sinema-ya   gid-er-iz
        we   every   week   movies-DAT   go-AOR-1PL
        'We go to the movies every week.'

                                      (Şener and Takahashi 2010)

(14) Bangla:
    a. Ami   æk-ṭa   boi-ke   poṛ-l-am
        1.SG   one-CL   book-ACC   read.PAST.1.SG
        'I have seen two books.'

    b. Tumi Ojothai cẽciye-mor-cho, keu   tomar   kOtha Sun-b-e na
        you   in.vain shout-die.PRES.2, nobody your   word hear-FUT.3 NEG
        'You are shouting in vain, nobody will hear you.'
                          (Bhattacharya 1997:12, Soma 2003:9)

(15) Ìgbo:
    a. Ọ   zà-bù-lù   ụnò
        3.SG   sweep.3.SG-PRI-AFF house
        'He/She swept the house.'

    b. Ị   zū-ọ-la   àlà³
        2.SG   buy.2.SG-OVS-PERF land³
        'You have bought land.'
                                      (Obiamalu 2015:53-80)

In all the examples above, only subjects show φ-agreement with verbal elements. If the lack of subject φ-agreement is a necessary condition for the productive Lexical VV-compounding, these languages should not allow such compounds. This prediction is not borne out. All these languages have a variety of Lexical VV-compounds as illustrated in (16) through (18).

---

³ OVS stands for the Open Vowel Suffix. See Obiamalu (2015:79) for details.

(16) Turkish:
    a. gelince-şaşır   b. geçip-git
       come-surprise   pass-go
       'come to surprise'   'leave through'
                                                                    (Kuribayashi 2006)

(17) Bangla:
    a. uRe-gE   b. ghumiye-poR
       fly-go   sleep-fall
       'fly away'   'fall asleep'
                                                                         (Soma 2003)

(18) Ìgbo:
    a. kụ́-wá   b. kwá-dà
       hit-break   push-fall
       'break by hitting'   'push (someone) down'
                                                                       (Ihionu 1992)

Given the observations above, I argue that it is the presence of object verb φ-agreement that makes productive VV compounding impossible in a language. The descriptive generalization in (12) is to be revised as in (19) below.

(19) Generalization (revised from (12))
      If a language lacks object-verb φ-agreement, then it can have Lexical VV-compounds.

In this section, we have observed data from six different languages; Korean, Mongolian, Malayalam, Turkish, Bangla, and Ìgbo. We reached a descriptive generalization of the relation between Lexical VV-compounds and φ-feature agreement. Only languages with subject φ-agreement or those without φ-agreement allow productive Lexical VV-compounds, while languages with both subject and object φ-agreement do not. It leads us to conclude that the presence of object-verb φ-agreement somehow prevents a language from forming Lexical VV-compounds. Then, the next question is, how we can most naturally explain this correlation syntactically, which is to be discussed extensively in the next section.

2. Minimal Search and the 'unambiguous path' in the vP-internal syntax

In order to formalize the descriptive generalization in (19), this section aims to propose a syntactic mechanism that blocks a derivation of Lexical VV-compounds. A non-phasal V inherits agreement inducing features from the phasal v, as in (20). Although C-to-T Feature Inheritance is optional in some languages, in which comlementizer agreement is observed (Haegeman 1992 among others), *v*-to-V feature inheritance is obligatory (see Chomsky 2008:149 for the discussions).
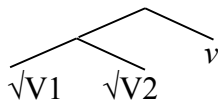
(20) Obligatory v-to-V Inheritance:
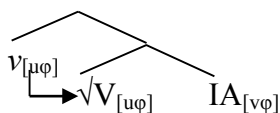
As for the structure of Lexical VV-compounds, I follow a slightly modified version of Nishiyama and Ogawa (2014), in which they are base-generated via set-merger.[4] Thus, VV compounds are formed when √V1 and √V2 merge, which is immediately dominated by *v*, as in (21).

(21) a. Merge(V1, V2) = {V1, V2}

√V1   √V2
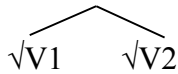
b. Introduce *v* (order irrelevant)

√V1   √V2   *v*

Why is VV-compounding blocked in languages with object-verb φ-agreement? I propose that it is the failure of Feature Inheritance that makes derivations unable to proceed. Since Agreement is a one-to-one relation,[5] I assume that Feature Inheritance, which is a prerequisite for Agreement, is also a one-to-one relation. Merger of √V1 and √V2 in (22) makes Feature Inheritance of [uφ] from *v* to √V becomes ambiguous since there are two √Vs.

(22) *Minimal Search cannot unambiguously relate v with φ-features to √V

$v_{[u\varphi]}$ …

√V1   √V2
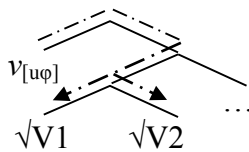
It makes Minimal Search unable to unambiguously relate *v* to √V, hence the derivation cannot proceed. Since Agreement cannot be induced, [uφ] on *v*, and possibly [uCase] and other [uF]s on the internal argument are left unvalued, which makes the Syntactic Object crash at the interfaces,[6] and the structure is deemed ungrammatical.

Some may say that Minimal Search may reach either √V1 or √V2 as a 'last resort' in order to uphold the derivation. Even if such Feature Inheritance is possible, I argue that the syntactic object results in ungrammaticality. Let us observe (23), in which [uφ] is on √V (either √V1 or √V2): The unvalued φ becomes unable to probe into [vφ] on the internal argument, as illustrated in (23), since [vφ] is no longer in its search domain (Note that √V is the probe, not *v* after Feature Inheritance).
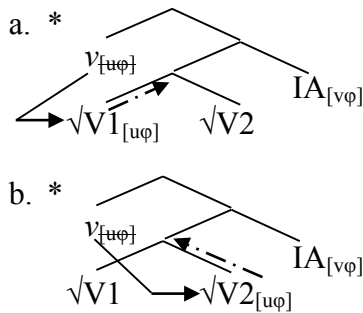
---

[4] In this paper, I do not assume pair-merge, pursuing the Simplest Merge conception (Chomsky 2013). I admit that many important issues are left open at this juncture, regarding pair-merge/adjunction and VV-compounds, which I continue to explore in my ongoing research.
[5] I do not assume multiple Agree of Hiraiwa (2005) here.
[6] In connection to this, some scholars assume that uninterpretable features may survive to the interfaces (Preminger 2014), though I do not discuss them in this article. Instead, I continue to follow Chomsky's (2000, 2008) and many others' assumptions that [uF] must be checked before they are transferred to the interfaces.

(23)   Probe[uφ] is unable to find its Goal[vφ]:

(order irrelevant)

a. *

$v_{[u\varphi]}$ $\sqrt{V1}_{[u\varphi]}$ $\sqrt{V2}$ IA$_{[v\varphi]}$

b. *

$v_{[u\varphi]}$ $\sqrt{V1}$ $\sqrt{V2}_{[u\varphi]}$ IA$_{[v\varphi]}$

As for *v* without unvalued φ-features, such as unaccusatives, I assume that *v* still has φ-features with some default values; hence Feature Inheritance obligatorily occurs, which results in the stoppage of derivation since Minimal Search cannot unambiguously relate *v* to √V in such a case as well as in (22). The abovementioned problems do not occur in languages without φ-features on *v*, since no Agreement or Feature Inheritance is obligatorily induced, as in (24).

(24)   VV-Compounding in languages with φ-featureless v:

v $\sqrt{V1}$ $\sqrt{V2}$ ...
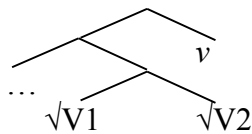
No φ-agreement/No Inheritance

So far, I have argued that the presence of φ-features on *v* makes Minimal Search unable to unambiguously relate it to multiple √Vs, which ends in ungrammaticality due to the halting of the derivation. The crucial support for my proposal is from another language in Bantu, Chicheŵa, in which object φ-agreement is optional (Bresnan and Mchombo 1987)[7]. The sentences in (25) are both grammatical, regardless of whether the object undergoes agreement with the verb or not.

(25)   a. Njûchi   ziná-lúm-a   alenje
          bees   3.PL.SUBJ-bite-PAST   hunters
          'The bees bit the hunters.'

       b. Njûchi   ziná-wa-lúm-a   alenje
          bees   3.PL.SUBJ-3.PL.OBJ-bite-PAST   hunters
          'The bees bit the hunters.'

(Adapted from Bresnan and Mchombo 1987)

If the object-verb φ-agreement blocks combining lexical verbs, then it predicts that lexical VV-compounds cannot be formed in Chicheŵa when the verb and object agree. The prediction is indeed borne out, as in (26) and (27). Although the VV-compound *ka-pemp* 'go-beg'

---

[7] I assume that *v* lacks [uφ], rather than having φ-features with default values when the object and verb show φ-agreement in Chicheŵa.

and *dza-man* 'come-refuse' can be formed in Chicheŵa (Baker 1988), they become unacceptable with the presence of object agreement φ-morphemes (Sam Mchombo p.c.).

(26) a. Ndi-ka-pemp-a   pamanga
       1.SG-go-beg-ASP   maize.3.PL
       'I go to beg maize.'

     b. *Ndi-wá-ka-pemp-a   pamanga
        1.SG.SUBJ-3.PL.OBJ-go-beg-ASP   maize.3.PL

     c. *Ndi-ka-pemp-wá-a   pamanga
        1.SG.SUBJ-go-beg-3.PL.OBJ-ASP   maize.3.PL

(27) a. Kati   madzi   banu   dza-man-e-ni   ine
       if   water   your   come-refuse-ASP-IMP   1.SG
       'If it is your water, come to refuse me.'

     b. *Kati madzi   banu   ndi-dza-man-e-ni   ine
        if   water   your   1.SG.OBJ-come-refuse-ASP-IMP   1.SG

     c. *Kati madzi   banu   dza-man-ndi-e-ni   ine
        if   water   your   come-refuse-1.SG.OBJ-ASP-IMP   1.SG

From these observations, I conclude that it is the presence of φ-features on *v* in the Lexicon that induces failure of Feature Inheritance and Agreement, which results in ungrammaticality of VV-compounds. The current analysis is compatible with the descriptive generalization in (19), which is repeated here as (28).

(28)   Generalization
       If a language lacks object-verb φ-agreement, then it can have Lexical VV-compounds.

Furthermore, the proposed analysis correctly explains why Lexical VV-compounding is possible in Chicheŵa only when the object and verb undergo φ-agreement. Although the proposed syntactic mechanism seems to capture the generalization in (19), the immediate question that arises is why English does not allow productive Lexical VV-compounds, though it does not show any overt evidence on *v* (i.e. φ-morphemes) for object-verb Agreement. Moreover, why is it the case that languages like Turkish, Bangla and Ìgbo do lack the object-verb Agreement, whereas English does have such Agreement? In the following section, I show that English does have indirect evidence for the object-verb φ-agreement, while Turkish, Bangla and Ìgbo do not.

3 Abstract Case vs. morphological case
3.1 Abstract Case valuation as a reflex of φ-agreement

English does not show any overt φ-morphology on verbs in object-verb agreement, as in (29). Nevertheless, as already noted, VV-compounds in English are not productive and a small

number of the examples such as *stir-fry*, *sleep-walk*, *slam-dunk*, et cetera. is frozen expressions (Lieber 2005:378).

(29)   a. The teacher scolded John/Mary/him/her/them/us.

         b. The student loves John/Mary/him/her/them/us.

At first sight, there is no evidence for object-verb φ-agreement in English. However, I argue that *v* has φ-features and undergoes Agreement with the object. Essentially along the lines of Chomsky (2000, 2008), I assume that Case features in languages such as English are valued as a reflex of φ-feature Agreement. That is, *v* licenses accusative Case, entering into Agree with the internal argument. An object's [uCase] is valued only when it values a φ-complete set of checking features on *v*. Therefore, although there is no *direct* evidence on verbs, DPs in English do have φ-features; hence an Agree relation with the DP has to involve φ-features, which is *indirect* evidence for the presence of object-verb φ-agreement. Due to the presence of φ-features on *v*, the blocking mechanism in section 2 rules out productive VV-compounding process in English.

So far, I have implicitly assumed that Turkish, Bangla and Ìgbo lack object verb φ-agreement. If the above discussion on English holds true in other languages as well, then any exception to the generalization in (28) may be considered to have covert object-verb φ-agreement, which makes the generalization vacuous or at least highly speculative.[8] In order to defend the claim that Turkish, Bangla, Ìgbo, Mongolian, Malayalam, Japanese and Korean lack object φ-agreement, this section is devoted to distinguishing the Case systems into two different types, namely Abstract Case and Morphological lower-case case systems (Kuroda 1988, Marantz 1992, McFadden 2004, Legate 2008 among others). The point here is that these languages entertain the latter morphological case system that makes no recourse to φ-agreement. For Japanese, many scholars have proposed different theories on case-licensing: Kuroda (1988), Marantz (1992), Aoyagi (1998), Fukui and Takano (1998), Saito (2012) inter alia, are each different in details but they agree that case-feature and φ-feature valuations are separated in Japanese unlike in English. Among them, this study specifically assumes Fukui and Takano's (1998) analysis, which claims that an argument's case is licensed by an overt presence of the morphological case head, K^0 (Bittner and Hale 1996).

3.2 In defense of the morphological case licensing

Let us now review some of the observations we made in section 2. The languages in question all have one feature in common: That is, they have overt case-particles attached to arguments. Japanese, Korean, Mongolian, Malayalam, Turkish, and Bangla in (30) through (35) all show overt morphological case-markers on the objects. As for Ìgbo, although it is not shown in the gloss, it has been observed that different tones are used to indicate cases on arguments (Echeruo 1998, Obiamalu 2015). Thus, all the languages listed below, which allow productive Lexical VV-compounds, have an overt morphophonological case system.

---

[8] I thank the anonymous reviewer of ConSOLE Proceedings for pointing this out.

(30)   Japanese and Korean:
       a. Taro-ga   Hanako-o   aisitei-ru.   (Japanese)
          T.-NOM   H.-ACC   love-PRES
          'Taro loves Hanako.'

       b. Mary-ka   chopap-ul   mek-ess-ta.   (Korean)
          M.-NOM   sushi-ACC   eat-PAST-DECL
          'Mary ate sushi.'

                                                                               (Aoyagi 2005:5)

(31)   Mongolian:
       a. Bat-Ø   nama-ig/chama-ig/ter-ig/bid-nig/tanar-ig/ted-nig   har-san.
          B.-NOM   me/you/him/us/you/them-ACC   see-PERF
          'Bat saw me/you/him/us/you/them.'

       b. Bi/Chi/Ter/Bid/Tanar/Ted   Bat-ig   har-san.
          I/You/He/We/You/They   B.-ACC   see-PERF
          'I/You/He/We/You/They saw Bat.'

                                                                               (Sakamoto 2011:33)

(32)   Malayalam:
       a. Siita   raaman-e   sneehiḱḱunnnu
          Sita   Raman-ACC   love
          'Sita loves Raman.'

       b. raaman   awan-te   bhaarya-ye   sneehiḱḱunnnu
          Raman   he-GEN   wife-ACC   love
          'Raman loves his wife.'

                                                                               (Jayaseelan 1999:30-44)

(33)   Turkish:
       a. Ben   bu   makale-yi   yavaş-yavaş   oku-yacağ-ım
          1.SG   this article-ACC   slowly-slowly   read-FUT-1.SG
          'I will read the article slowly.'

       b. Biz   her   hafta   sinema-ya   gid-er-iz
          we   every   week   movies-DAT   go-AOR-1.PL
          'We go to the movies every week.'

                                                                               (Şener and Takahashi 2010:86)

(34) Bangla:
  a. Ami   æk-ṭa   boi-ke   poṛ-l-am
     1.SG   one-CL   book-ACC   read.PAST.1SG
     'I have seen two books.'

  b. Tumi Ojothai   cĕciye-mor-cho, keu   tomar   kOtha Sun-b-e   na
     you   in.vain   shout-die.PRES.2, nobody your   word   hear-FUT.3 NEG
     'You are shouting in vain, nobody will hear you.'

                                                      (Bhattacharya 1997:12, Soma 2003:9)

(35) Ìgbo:
  a. Ọ   zà-bù-lù   ụnọ̀
     3.SG   sweep.3.SG-PR-AFF house
     'He/She swept the house.'

  b. Ị   zū̄-ọ-la   àlà
     2.SG   buy.2.SG-OVS-PERF   land
     'You have bought land.'

                                                              (Obiamalu 2015:53-80)

I argue that the existence of overt case particles/tones case licenses objects with no recourse to φ-agreement in these languages. Specifically, I assume Fukui and Takano's (1998:58) KP-analysis, which follows Kuroda's (1988) morphological case licensing system (the lower-case case licensing) *vis-à-vis* abstract Case licensing.

(36) Principle of Morphological Licensing:
      Arguments must be licensed by Case (abstract) or case (morphological)
                                                                    (Kuroda 1988:40)

They assume that there are two types of Case/case licensing systems provided in UG, the latter of which is entertained by languages such as Japanese. In order to provide an explicit account of how accusative is licensed in these languages, I assume with Fukui and Takano that a case particle is a morphological realization of a case feature and heads its own projection, KP (Bittner and Hale 1996), dominating a nominal phrase, as illustrated in (37).[9]

(37) a. John-ga   Mary-o   mi-ta
         J.-NOM   M.-ACC   see-PAST
         'John saw Mary.'

      b.   KP
          ⟋   ⟍
      Mary   -o (=K$_{ACC}$)

                                                          (Fukui and Takano 1998:55)

With no recourse to feature checking via φ-agreement, nominals are licensed by the morphologically overt case particle, K in (37). In Fukui and Takano's (1998) system, the accusative

---

[9] Since Nominative Case assignment to subjects in these languages does not affect the discussions here, I leave it open for future research. See Fukui and Takano (1998) for details.

case assignment feature is linked to a particular θ-role on the transitive verb, which is summarized in (38).

(38)   a. The Case particle makes the Case feature of a noun phrase visible to Spell-Out.

       b. Linking to a particular θ-role makes the Case feature of a (transitive) verb interpretable at LF.

                                        (Adapted from Fukui and Takano 1998:59)[10,11]

Following the spirit of Fukui and Takano, I argue that the NP is case-licensed if K_{ACC} appears on the NP that receives an object-related θ-roles (e.g. $θ_{theme}$), as in (39).[12] Thus, the present analysis claims that in languages with a rich system of overt morphophonological case particle/tones, the accusative is more like an inherent Case, rather than a structural Case (Takahashi 1993, Fukui and Takano 1998 inter alia).

(39)   KP is licensed via merger with V (i.e. the √ROOT categorized by v):
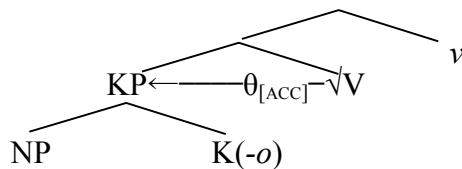
A language can have both abstract Case licensing and morphological case licensing systems such as Turkish, Bangla and Ìgbo in (33) through (35). If one takes the Kuroda's (1988) or Fukui and Takano's (1998) morphological case system, the accusative assignment in languages with rich case morphology need not make recourse to object φ-agreement, which is compatible with the generalization in (28), repeated here as (40).

(40)   Generalization
       If a language lacks object-verb φ-agreement, then it can have Lexical VV-compounds.

---

[10] A reinterpretation of Fukui and Takano's KP-analysis is also made by Zushi (2016):
    The mechanism could be reformulated as a case licensing. We could argue that a nominal expression is merged with a case marker to form KP, and that the KP is case licensed when it is merged with a particular element. For example, a nominal combined with an accusative marker (ACCP) is licensed when it is merged with a verb. (Zushi 2016:62)
[11] This is compatible with Zushi's (2016) proposal that case in Japanese is valued via merger of nominals and certain kinds of phasal/non-phasal heads as in (i) through (iii).
(i)   a. When a nominal is merged with a lexical head, its case feature is valued as accusative.
        b. When a nominal is merged with a phrase head (v, or n), its case feature is valued as nominative or genitive.
        c. Otherwise, the case feature of a nominal is valued as dative. (Zushi 2016:48)
[12] I do not consider θ-roles to be features. Regarding predicate-argument relations, I assume with Hale and Keyser (1993), Grimshaw (1990), Larson (1988) and many others that they are configurationally determined between the predicate and its arguments in Narrow Syntax. How thematic roles are assigned to arguments raises the so-called *linking problem* (Baker 1997). Baker (1997:126) states that thematic relations may be nothing other than the way SEM "reads off" syntactic relations such as the specifier of VP or complement of V. I basically follow this view, which is rooted in Hale and Kayser (1993) and Reinhart's series of works (c.f. Reinhart 2002): The so-called *thematic roles* are "just the relations determined by the categories and their projections" (Hale and Kayser 1993:53): "[T]he thematic terminology typically applied in this case simply reflects the relational status of the NP" (Hale and Kayser 1993:68).

4. Conclusion

In this paper, I highlighted a correlation between the absence of object-verb φ-agreement and the presence of productive Lexical VV-compounds. Furthermore, I proposed that Feature Inheritance as a prerequisite for object-verb φ-agreement makes Minimal Search unable to unambiguously relate v to $\sqrt{}$V, which results in the halting of the derivation. It deems the structure ungrammatical in languages with φ-features on v. This is exactly why certain languages lack productive Lexical VV-compounds, while languages such as Japanese abound in it.

The proposal in this paper is compatible with the fact that Germanic languages like English, German and Dutch (Booij 1992) allow productive root compounds of categories other than VV, such NN, NA, AA, AN, etc. (Roeper et al 2002), as in English examples in (41).

(41) a. NN: *dog-bowl, apple-cake, ice-cream, file-cabinet, towel-rack, catfood, steelmill, …*
     b. NA: *sky-blue, leaf-green, stone-cold, rock-hard, …*
     c. AA: *red-hot, blue-green (green-blue), icy-cold, wide-awake, …*
     d. AN: *blackboard, dry-farm, wet-sand, double-coat, sweet-talk, …*
                         (Lieber 2005:378 (from COCA and other Corpora))

This analysis correctly predicts that such compounds are rich in these languages, whereas VV-compounds are frozen expressions and barely productive (Lieber 2005, Bauer et al. 2013). Verbs induce φ-agreement and the blocking mechanism in Section 2 precludes productive VV compounds. On the other hand, NN, NA, AA or AN are not restricted in such a way; hence productive root compounding of these non-verbal categories correctly receives a natural analysis as the syntactic merger of heads, as Roeper et al. (2002) claim.[13]

Before closing the article, some discussions are in order. In languages such as German, expressions such as *kennen lernern* 'get to know', *spazieren gehen* 'take a walk' and the like exist. At first glance, they seem to be the Lexical VV-compounds, and indeed Neef (2009) and others classify them as VV-compounds. However, I argue that they are not: They are called *Partikelverben*, which are separable verbs. Indeed, they are separated due to the verb-second phenomenon (Peter Erdmann, Haider Hubert p.c.), as in (42), since it violates the non-interruptibility principle of lexical integrity in the definition of the endocentric VV-compounds in (3), which is repeated here as (43).

(42) a. Ich   lerne   keinen Mann   kennen
         I   learn   no   man   know
         'I get to know nobody.'

     b. *Ich kennen-lerne   keinen Mann
         I   know-learn   no   man

(43) Endocentric VV-compounds that satisfy the non-interruptibility principle of lexical integrity, which behave as indivisible V0 units in the phrasal syntax.
                                        (Kageyama 2016:278)

The overall discussions pointed out the correlation between productive Lexical VV-compounds and the lack of φ-features on v and the object-verb φ-agreement in comparative perspectives. Thus, the current study supports the Fukui-Borer Thesis, or the Chomsky-Borer

---

[13] See Kiparsky (2009) for [$_V$ N+N] analysis of exocentric Greek dvandva compounds.

Conjecture: Cross-linguistic variation is limited to differences in the properties of certain functional elements (here, whether *v* has φ-features or not) in the Lexicon. Although this article has left open several important issues, I hope to have shown that the analysis proposed in this paper provides a promising insight into the nature of cross-linguistic variations of Lexical VV-compounds.

Ryoichiro Kobayashi
Graduate School of Languages and Linguistics,
Sophia University
Japan Society for the Promotion of Science
ryoichiro.k@eagle.sophia.ac.jp
https://sites.google.com/site/ryoichirokobayashi/

Abbreviations

NOM	nominative	AFF	affirmative
ACC	accusative	NEG	negation
DAT	dative	SUBJ	subject
GEN	genitive	OBJ	object
AOR	aorist	CAUSE	causative
PRES	present	PASS	passive
PAST	past	SG	singular
FUT	future	PL	plural
ASP	aspect(ual)	1	first person
PERF	perfective	2	second person
IMP	imperfective	3	third person
PRI	prior	CL	classifier
DECL	declarative	OVS	open vowel suffix

References

Aoyagi, Hiroshi. 2006. On predicate focus constructions in Korean and Japanese. *Harvard Studies in Korean Linguistics*, 11, 359-372.

Aoyagi, Hiroshi. 2011. Morphological case vs Abstract Case: Toward an explanatory theory of Japanese case marking. Handout presented at the KGGC November Meeting.

Aoyagi, Hiroshi. 2014. On serialized verbs in Japanese and Korean. *Harvard Studies in Korean Linguistics* 15, 219-231.

Baker, Mark. 1988. *Incorporation*. Chicago: University of Chicago Press.

Baker, Mark. 1997. Thematic roles and syntactic structure. In Rizzi, L., & L. Haegeman. (eds.), *Elements of Grammar*, 73-137. Netherlands: Springer.

Bauer, Laurie., Lieber, Rochelle., and Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Bittner, Maria., and Kenneth Hale. 1996. The structural determination of case and agreement. *Linguistic inquiry*, 27(1), 1-68.

Booij, Geert. 1992. Compounding in Dutch. *Rivista di Linguistica* 4(1), 37-60.

Bresnan, Joan., and Sam Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language,* 741-782.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Martin, R. et al (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Freidin, R. et al. (eds.), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 291-321. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua*, 130, 33-49.

David, Anne, Boyle. 2015. *Descriptive Grammar of Bangla*. Berlin: Walter de Gruyter.

Echeruo, Michael. 1998. The case-tone factor in Ìgbo nouns with special reference to the Ìgbo associative construction. *Language History & Linguistic Description in Africa*, 147-153.

Fukuda, Shin. 2012. Aspectual verbs as functional heads; evidence from Japanese aspectual verbs. *Natural Language & Linguistic Theory*, 30(4), 965-1026.

Fukui, Naoki. 1986. *A theory of category projection and its applications*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Fukui, Naoki. 1988. Deriving the differences between English and Japanese: A case study in parametric syntax. *English Linguistics*, *5*, 249-270.

Fukui, Naoki. 1995. Approach: A comparative syntax of English and Japanese. In M. Shibatani and T. Bynon (eds.), *Approaches to language typology*, 327-372. Oxford: Oxford University Press.

Fukui, Naoki., and Yuji Takano. 1998. Symmetry in syntax: Merge and demerge. *Journal of East Asian Linguistics*, 7(1), 27-86.

Fukushima, Kazuhiko. 2005. Lexical VV compounds in Japanese: Lexicon vs. syntax. *Language*, 568-612.

Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, Mass: MIT Press.

Haegeman, Liliane. 1992. *Theory and description in generative syntax, A case study in West Flemish*. Cambridge: Cambridge University Press.

Hale, Kenneth., and Samuel J Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Hale, K., & S. J. Keyser. (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 53-109, Cambridge, MA: MIT Press.

Hiraiwa, Ken. 2005. *Dimensions of symmetry in syntax: Agreement and clausal architecture.* Cambridge, MA: Massachusetts Institute of Technology dissertation.

Ihionu, Peter. 1992. Verb compounding in Igbo: an overview. In Collins, C., & V. Manfredi (eds.), *Proceedings of the Kwa Comparative Syntax Workshop*, *MIT Working Papers in Linguistics*, 17, 165-182.

Jayaseelan, K. A. 1999. *Parametric studies in Malayalam syntax*. New Delhi: Allied Publishers.

Kageyama, Taro. 1993. *Bunpoo to Gokeesei* [*Grammar and Word Formation*]. Tokyo: Hitsuji Shobo.

Kageyama, Taro. 1999. Word Formation. In Tsujimura, N. (ed.). *The Handbook of Japanese Linguistics*, 297-325. Malden, MA: Blackwell Publishers Ltd.

Kageyama, Taro. 2016. Verb-compounding and verb-incorporation. In Kageyama, T., & H. Kishimoto (eds.). *Handbook of Japanese Lexicon and Word Formation*, 273-310. Berlin: Walter de Gruyter GmbH & Co KG.

Khurelbat, Bandzragchin. 1992. *Word formation in Mongolian language.* New Delhi: Jawaharlal Nehru University dissertation.

Kiparsky, Paul. 2009. Verbal co-compounds and subcompounds in Greek. In Halpert, C. et al (eds.), *Proceedings of the 2007 Workshop in Greek Syntax and Semantics*. *MIT Working Papers in Linguistics 57*, 187-195. Cambridge MA: MIT Press.

Ko, Heejeong., and Daeyoung Sohn. 2015. Decomposing complex serialization: The role of v. *Korean Linguistics*, 17(1), 78-125.

Kobayashi, Ryoichiro. 2016. φ-defectiveness and the predicate-internal Syntax: Comparative syntax of VV-compounds and their kin. Paper presented at the workshop "Current Issues in comparative Minimalist syn-

tax" at the 34th annual meeting of the English Linguistic Society of Japan, November 12th, Kanazawa University.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Mass: MIT Press.

Kuribayashi, Yu. 2006. Torukogo-no Hukugoodooshi to Bumpooka. [On the V-V Compound in Turkish and its Grammaticalization]. *Azia, Ahurika-no Gengo to Gengogaku*, 1, 25-44.

Kuroda, Shige-Yuki. 1988. Whether we agree or not: A comparative syntax of English and Japanese. *Lingvisticae Investigationes*, 12(1), 1-47.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry*, 19(3), 335-391.

Legate, Julie. 2008. Morphological and abstract case. *Linguistic Inquiry*, 39(1), 55-101.

Lieber, Rochelle. 1992. Compounding in English. *Rivista di linguistica*, 4(1), 79-96.

Lieber, Rochelle. 2005. English word-formation processes. In Štekauer, P., & R. Lieber. (eds.). *Handbook of word-formation*, 375-427. Netherlands: Springer.

Marantz, Alec. 1992. Case and licensing. In German, W. et al. (eds.), Proceedings of the Eastern States Conference on Linguistics, 234-252. Columbus: Ohio State University.

McFadden, Thomas. 2004. *The position of morphological case in the derivation*. Philadelphia, PA: University of Pennsylvania dissertation.

Moravcsik, Edith. 1974. Object-verb agreement. *Working Papers on Language Universals*, 15, 25-140.

Neef, Martin. 2009. IE, Germanic: German. In Lieber, R., & P. Štekauer. (eds.). *The Oxford handbook of compounding*, 386-399. Oxford: Oxford University Press.

Nishiyama, Kunio. 1998. VV compounds as serialization. *Journal of East Asian Linguistics,* 7(3), 175-217.

Nishiyama, Kunio. 2008. V-V compounds. In Miyagawa, S., & M. Saito. (eds.). *Handbook of Japanese linguistics*, 320-347, New York: Oxford University Press.

Nishiyama, Kunio., and Yoshiki Ogawa. 2014. Auxiliation, atransitivity, and transitivity harmony in Japanese VV compounds. *Interdisciplinary Information Sciences*, *20*(2), 71-101.

Obiamalu, Greg Orji. 2014. Negation and V-movement in Igbo: a minimalist perspective. *SKASE Journal of Theoretical Linguistics*, 11(2), 42-59.

Obiamalu, Greg Orji. 2015. Functional Categories in Igbo: A minimalist perspective. Port Harcourt: M & J Grand Orbit Communications.

Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.

Roeper, Thomas., Snyder, William., and Kazuko Hiramatsu. 2002. Learnability in a minimalist framework: Root compounds, merger, and the syntax-morphology interface. *The process of language acquisition*, 25-35.

Saito, Mamoru. 2012. Case checking/valuation in Japanese: Move, agree, or merge. *Nanzan linguistics*, 8, 109-127.

Sakamoto, Yuta. 2011. *A study of sluicing and cleft in Mongolian: A comparison with Japanese.* MA thesis, Tohoku University.

Şener, Serkan., and Daiko Takahashi. 2010. Ellipsis of arguments in Japanese and Turkish. *Nanzan Linguistics*, *6*, 79-99.

Soma, Paul. 2003. Composition of compound verbs in Bangla. In Beermann, D., & L. Hellan (eds.), *Proceedings of the workshop on multi-verb constructions*, 1-18. Trondheim: Norwegian University of Science and Technology.

Takashashi, Daiko. 1993. On antecedent contained deletion, ms., University of Connecticut.

Washio, Ryuichi. 1995. *Interpreting Voice: A case study in lexical semantics*. Tokyo: Kaitakusha.

Yashima, Jun. 2008. *Clausal architecture and complex predicate formation: A case study of Japanese syntactic VV compounds*. Tokyo: University of Tokyo MA thesis.

Yumoto, Yoko. 2005. *Fukugodoshi haseidoshi no imi to togo* [*Semantics and syntax of compound verbs and complex verbs*]. Tokyo: Hituzi Shobo.

Zushi, Mihoko. 2016. Case and predicate-argument relations. In Fujita, K., & C. Boeckx (eds.), *Advances in Biolinguistics: The Human Language Faculty and Its Biological Basis*, 46-66. London and New York: Routledge.

A configurational account of Turkish differential subject marking

Sabine Laszakovits

Turkic languages have been argued to require a configurational analysis for the distribution of structural accusative and dative, but it is debated whether this should extend to nominative and genitive. This paper argues that a configurational analysis can be upheld for nominative/genitive alternations in differential subject marking (DSM) constructions by analyzing argument clauses, but not adjunct clauses, as complex NPs with an optionally null head noun. This head noun creates the nominal case assignment domain required for the assignment of genitive case.

1.   Introduction

Differential case marking refers to the property of some languages to mark the same grammatical function in different morphological ways, often depending on factors such as definiteness or animacy of the respective argument (Bossong 1985). Turkish has at least three kinds of differential case marking: 1. differential object marking (DOM), which marks only definite or specific direct objects with accusative, 2. differential subject marking (DSM) in nominalized clauses, which marks only definite and specific subjects with genitive, and 3. DSM in nominalized clauses, which blocks genitive-marking according to 2. on subjects of adjunct clauses, but not on subjects of argument clauses, the complements of complex NPs, or of relative clauses.

The distribution of DOM, among many other properties of accusative and structural dative, follows straight-forwardly from a configurational approach to case-assignment (Baker & Vinokurova 2010, Baker 2015). The goal of this paper is to extend the analysis of Turkish cases as configurational cases to nominative and genitive. To do this, we will look more closely at the environment that licenses genitive case in nominalized clauses, and why this licensing does not happen in adjunct clauses.[1]

---

[1]Other environments for genitive assignment are i. the modifier in a possessive construction (e.g. _my book_), ii. the subject complement in nominal sentences (e.g. _This book is mine_), and iii. the subject complement in small clauses (e.g. _John considers this book (to be) mine_) (Göksel & Kerslake 2005:§14.3.3.6, pp. 181f.). The environments for nominative case are harder to define because nominative is morphologically null and therefore does not

The structure of this paper is as follows: in §1.1, I will introduce the basic paradigm of four types of embedded nominalized clauses and their case marking properties. Section 2.1 illustrates existing evidence in favor of a dependent case approach for Turkic, and §2.2 presents the present proposal. In §3, I discuss predictions of my analysis and show that they are borne out. Section 4 concludes.

1.1. Turkish DSM based on clause type

Turkish has several types of nominalized clauses. They share the property that verbal agreement appears in the possessive paradigm rather than the personal paradigm, and that the nominalized verb can receive case endings in accordance with its function in the matrix clause. This paper will only be interested in the factive nominalizers *-DIG* and *-(y)AcAG* because they are the only nominalizers that exhibit the full paradigm of subject marking alternations, as we will see shortly. To my knowledge, they only differ in temporal orientation: *-DIG* is used for past and present tense, while *-(y)AcAG* is used for future tense.

Of interest to us are four different types of clauses marked with *-DIG*: argument clauses, adjunct clauses, non-subject relative clauses, and clausal complements in complex NPs.

First, we see an argument clause with the nominalizer *-DIG* in (1). In (1a) the argument clause is marked with structural accusative, in (1b) with lexical ablative, and in (1c) with nominative. The internal structure is the same in all three sentences, and the subject bears genitive case.

(1)   a.   [Sen-*(in)   dün   sabah   ev-de   yemek pişir-diğ-in   ]-i
            you-GEN/*NOM yesterday morning home-LOC food   cook-*DIG*-2SG -ACC
            duy-du-m.
            hear-PST-1SG
            'I heard that you had cooked food at home yesterday morning.' (Kornfilt 2003:(6))

    b.   [Sen-*(in)   yemek pişir-diğ-in   ]-den kork-tu-m.
            you-GEN/*NOM food   cook-*DIG*-2SG -ABL fear-PST-1SG
            'I feared that you had cooked food.'

    c.   [Sen-*(in)   yemek pişir-diğ-in   ]-∅   bana   sürpriz yap-tı.
            you-GEN/*NOM food   cook-*DIG*-2SG -NOM me.DAT surprise do-PST
            'That you cooked food surprised me.'

In argument clauses the subject has the option to lose its genitive marking when it is adjacent to the verb, indefinite, and not specified for number, as shown in (2a) with an indefinite subject and no genitive versus (2b) with a definite subject and genitive. This type of DSM can be explained by "pseudo-incorporation" of the subject into the predicate in order to avoid the Case Filter (e.g. Kornfilt 1997b:396–400) as is commonly assumed for the parallel DOM data (Nilsson 1986, Enç 1991, Aygen 1999, Aydemir 2004, von Heusinger & Kornfilt 2005; i.a., see also Baker & Vinokurova 2010 for Sakha).

---

contrast on the surface with the absence of case (Göksel & Kerslake 2005:§14.3.3.1, pp. 173–5; see also Kornfilt & Preminger 2015 on the idea that nominative *is* no case at all).

(2)   a.   [Çocuğ-u   arı-∅   sok-tuğ-un   ]-u   duy-du-m.
            child-ACC bee-NOM sting-*DIG*-3SG -ACC hear-PST-1SG
            'I heard that <u>bees</u> stung the child.'

       b.   [Çocuğ-u   arı-nın   sok-tuğ-un   ]-u   duy-du-m.
            child-ACC bee-GEN sting-*DIG*-3SG -ACC hear-PST-1SG
            'I heard that <u>the bee</u> stung the child.'

Second, if the nominalized clause is not selected for by the matrix verb, i.e. if it is an adjunct clause, the subject cannot be marked with genitive case, instead it is obligatorily in the nominative, even when referential and not adjacent to the verb, as shown in (3).

(3)   a.   Ben   [Ali-(*nin)   cam-ı   kır-dığ-ı   zaman] gerçeğ-i
            I.NOM   Ali-NOM/*GEN glass-ACC break-*DIG*-3SG when   truth-ACC
            bil-iyor-du-m.
            know-PROG-PST-1SG
            'I knew the truth when Ali broke the glass.'   (Aygen 2007:(3))

       b.   [Hasan-(*ın)   söylenti-yi duy-duğ-un-a   göre   ] herkes
            Hasan-NOM/*GEN rumor-ACC hear-*DIG*-3SG-DAT because   everybody
            duy-acak.
            hear-FUT
            'Since Hasan heard the rumor, everybody will hear it.'

Following Aygen (2007), I analyze *zaman* and *göre* in (3) as complementizers that historically derive from nouns (*zaman* 'time') or postpositions (*göre*+DAT 'according to'). Evidence for this comes from the fact that synonyms of *zaman* cannot appear as the head of adjunct clauses: replacing *zaman* with *vakit* 'time' in (4) makes the sentence ungrammatical.

(4)   *Ben   [Ali-(nin)   cam-ı   kır-dığ-ı   vakit] gerçeğ-i
           I.NOM   Ali-NOM/GEN glass-ACC break-*DIG*-3SG time   truth-ACC
           bil-iyor-du-m.
           know-PROG-PST-1SG

Further evidence comes from modification: while the noun *zaman* in the relative clause in (5a) can be modified with the plural suffix, this is not possible in (5b) where *zaman* is not an argument in the matrix clause:

(5)   a.   [[Mehmet-le   kavga yap-tığ-ımız] zaman-(lar)]-∅   zor-du.
            Mehmet-with quarrel do-*DIG*-1PL   time-PL   -NOM hard-PST
            'The times when we used to fight with Mehmet were hard.'

       b.   [Mehmet-le   kavga yap-tığ-ımız zaman-(*lar)] evlen-di-m.
            Mehmet-with quarrel do-*DIG*-1PL time-PL   marry-PST-1SG
            'I got married when we were fighting with Mehmet.'

Third, nonsubject relative clauses like (5a) also use nominalized *-DIG* or *-(y)AcAG* clauses, and the subject is marked with genitive case. We can test for the gap by overtly realizing the relativized constituent, which is not possible:

(6)   [Ali-*(nin)   (*dün   5:32-de) cam-ı   kır-dığ-ı   zaman]-ı
          Ali-GEN/*NOM (yesterday 5:32-LOC) glass-ACC break-*DIG*-3SG time   -ACC

bil-iyor-du-m.
know-PROG-PST-1SG
'I knew when Ali broke the glass (*yesterday at 5:32).' (Aygen 2007:(2))

Fourth, the nominalized clausal complement in a complex NPs also marks its subject with genitive case:

(7) Ben [[Hasan-*(ın) Jale-yi gör-düğ-ü] gerçeğ-in]-i
I.NOM Hasan-GEN/*NOM Jale-ACC see-*DIG*-3SG fact-3SG -ACC
bil-iyor-um.
know-PROG-1SG
'I know the fact that Hasan saw Jale.' (Aygen 2007:(14))

Summarizing, in argument clauses, complex NPs, and relative clauses, the subject of a *-DIG/-(y)AcAG* clause is marked with genitive case, while in otherwise identical adjunct clauses it is marked with nominative case.

2. *The proposal*

In §2.1, I provide motivation for a dependent case approach by summarizing arguments that Turkic cases are best analyzed as configurational cases. In §2.2, I present my proposal regarding how to analyze the nominative/genitive alternation based on clause type in a configurational framework.

2.1. *Evidence for dependent case in Turkic*

A configurational system of case assignment, such as Dependent Case Theory (Yip et al. 1987, Marantz 1992, Bittner & Hale 1996), states that cases are assigned based on the configurations of one or more noun phrases in a case assignment domain — in particular depending on whether a given NP is c-commanded/c-commands another NP and what the domain is. Structural case marking serves to distinguish multiple noun phrases in the same case assignment domain from each other. When the grammar encounters two or more NPs that have not been marked with lexical cases, *dependent cases* are fired, which are specified for a case, a case assignment domain, and whether to mark the unmarked NP c-commanding other unmarked NPs, or the one being c-commanded by other unmarked NPs. After this step, any remaining unmarked NPs are marked with *unmarked cases*, which are specified for a case assignment domain.

This can be seen straightforwardly with accusative and dative case assignment in Turkic: Baker & Vinokurova (2010), Baker (2015) analyze Sakha accusative case as a dependent case that gets assigned to a c-commanded, unmarked NP in the CP-domain. This predicts that the direct object will only receive ACC if there is a higher unmarked NP, such as a subject, and if the DO leaves the vP-domain and moves into the CP-domain. This is consistent with the DOM patterns: only specific or definite objects receive ACC, and they must precede low adverbs, while unmarked direct objects are necessarily non-specific and indefinite and follow low adverbs. The facts for Turkish are identical and illustrated in (8).

(8)   a.   Ben   hızlı kitap-(*ı)   oku-r-um.
              I.NOM fast book-*ACC read-AOR-1SG
              'I read fast.' (lit: 'I'm fast at book-reading.')
        b.   Ben   kitab-*(ı) hızlı oku-r-um.
              I.NOM book-ACC fast read-AOR-1SG
              'I will read the book fast.'   (Aygen 1999:13f.)

Furthermore, Sakha and Turkish have so-called raising-to-object, whereby in finite embedded clauses the NOM-subject can raise to the edge of the embedded clause and receive ACC-marking (Şener 2008, 2011).

(9)   Pelin [ben-(i)   Timbuktu-ya git-ti-m   (diye)] sanı-yor.
        Pelin   I-NOM/ACC Timbuktu-DAT go-PST-1SG that   think-PROG
        'Pelin thinks that I went to Timbuktu.'   (Şener 2008:1f.)

Crucially, this ACC is only available if there is a higher unmarked NP in the matrix clause, and does not depend on the transitivity properties of matrix *v*, posing a problem for accounts of accusative checking via *v*. In (10), illustrated for Sakha, the matrix predicate 'become sad' is intransitive, yet accusative can be licensed. This contrasts with (11), where there is no matrix subject and accusative is not available.

(10)   Keskil [Aisen-y   kel-bet   dien] xomoj-do.
        Keskil   Aisen-ACC come-NEG.AOR that   become.sad-PST
        'Keskil became sad that Aisen is not coming.'

                                             (Sakha; Baker & Vinokurova 2010:617)

(11)   [Aisen-(*y)   massyyna atyylah-ar-a   ] naada buol-la.
          Aisen-*ACC car   buy-AOR-3SG   need become-PST
        'It became necessary for Aisen to buy a car.'   (Sakha; B&V:619)

This raising-to-object also allows the co-occurrence of two accusative-marked NPs in the same clause, as shown in (12). This is predicted under a dependent case analysis of accusative: 'house' receives ACC because it is c-commanded by case-less 'Masha', and the same applies to 'Misha'. In a second step, 'Masha' receives NOM as unmarked case.

(12)   Masha [Misha-ny kel-ie   dien] djie-ni   xomuj-da.
        Masha   Misha-ACC come-FUT that   house-ACC clean-PST
        'Masha tidied up the house (thinking) that Misha would come.'   (Sakha; B&V:618)

While accusative is the dependent-down case in the CP-domain, structural dative case is the dependent-up case in the VP-domain, i.e., it is assigned to the higher one of two case-less NPs. We see dative/accusative alternations in causatives, where the case of the causee depends on the valency of the predicate. In (13), causativizing an intransitive yields accusative-marking of the causee. In (14), causativizing a transitive produces dative-marking.

(13)   a.   Ali ağla-dı.
   Ali cry-PST
              'Ali cried.'

  b. Bahar Ali-yi ağla-t-tı.
    Bahar Ali-$\boxed{\text{ACC}}$ cry-CAUS-PST
    'Bahar made Ali cry.'

(14) a. Caner Deniz-i öp-tü.
    Caner Deniz-ACC kiss-PST
    'Caner kissed Deniz.'

  b. Emre Caner-e Deniz-i öp-tür-dü.
    Emre Caner-$\boxed{\text{DAT}}$ Deniz-ACC kiss-CAUS-PST
    'Emre made Caner kiss Deniz.'

When dependent case marking takes place in (14b), there are two case-less NPs inside the VP-domain, so the higher one ('Caner') gets marked with dative case. Then the lower one ('Deniz') moves into the CP-domain where it encounters the case-less subject and receives accusative-marking. (Then more scrambling happens.)

 The same procedure applies to goals of ditransitives.

 A configurational approach also correctly predicts that causees of ditransitives will receive dative marking, resulting in a pattern (NOM)–DAT–DAT–ACC: such a configuration is given in (15), where the following dependency relationships hold in the VP-domain: YP c-commands case-less ZP, so YP receives DAT. XP c-commands case-less ZP, so XP receives DAT. ZP does not receive any case in the VP-domain, but will receive accusative in the CP-domain if it moves there.

(15)

         VP
        ╱ ╲
      XP   ╲
       YP ╲
        ZP
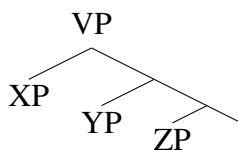
Having seen that a dependent case approach for Turkic accusative and dative not only fares very well but can explain data that assignment of accusative by *v* cannot (10), I now turn to my proposal of how to assign nominative and genitive case in this system.

2.2. *Genitive as unmarked case in a nominal environment*

I adopt Lees's (1965) and Aygen's (2007) idea that Turkish argument clauses are complex NPs with a silent abstract head noun — contrasting with adjunct clauses, which are CPs — and I argue that this derives the correct results in a configurational theory of case. Concretely, I follow Marantz (1992) in assuming that nominative and genitive case are both *unmarked* cases that are assigned in different case assignment domains: nominative is the unmarked case in a clausal environment (CP), and genitive case is the unmarked case in a nominal environment (NP). I adopt Baker's (2015) claim that case assignment domains correspond to the complements of phase heads. Phase extensions to the silent head noun will then create a nominal case assignment domain and result in genitive case on the embedded subject.

 The structure of argument clauses and complex NPs that I propose is given in (16): a full CP with an empty complementizer is dominated by an NP with an optionally overt or covert head noun. Embedding a full CP under the head noun is necessary because they can contain
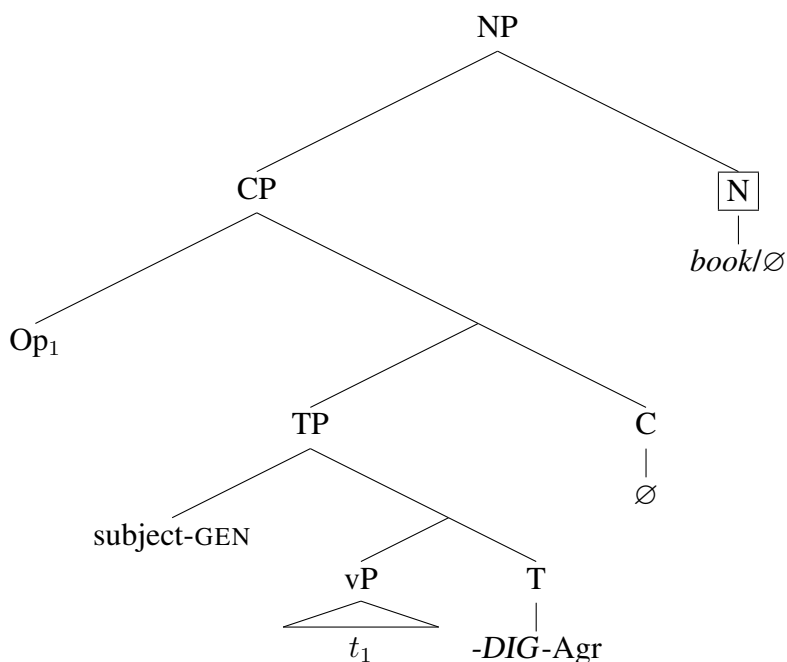
wh-words and mood expressions (Kornfilt 2006).

(16)   Argument clause / complex NP:

```
                        NP
                 _____/_____
                CP              N
          _____/\_____         |
         TP           C     rumor/∅
    ____/\____        |
subject-GEN           ∅
          __/\__
         vP      T
        /__\     |
             -DIG-Agr
```
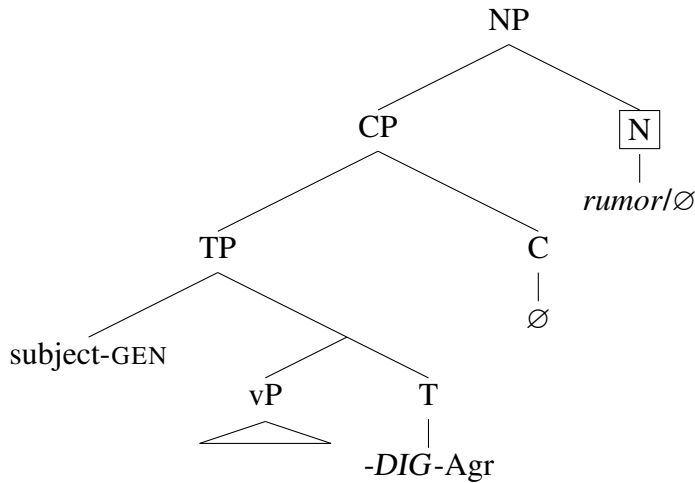
Genitive case is assigned to the subject in Spec,TP as unmarked case in a nominal case assignment domain. The relevant case assignment domain is triggered by the head noun (here: 'rumor'). Baker (2015) assumes as null hypothesis that case assignment domains correspond to spell-out domains, i.e. to the complements of phase heads. Having evidence that these clauses are CPs by virtue of them holding wh-elements and mood, and assuming that CPs are phases, I propose that the shift from a clausal domain to a nominal domain proceeds by phase extension (den Dikken 2007, Gallego 2010, Wurmbrand 2013; see also Colley & Davis to appear), i.e. incorporation of the null complementizer into the head noun, making N the phase head (phase heads are boxed in (16)–(18)).

The same happens in non-subject relative clauses, which I assume to merge an operator in Spec,CP (following Kornfilt 1997a).

(17)   Relative clause:

```
                          NP
                  _____/_____
                 CP                 N
          _____/_____          |
       Op_1            TP        book/∅
                  ____/\____
                 TP          C
            ____/\____       |
        subject-GEN          ∅
                  __/\__
                 vP      T
                /__\     |
                 t_1  -DIG-Agr
```

Adjunct clauses differ from argument clauses, complex NPs, and relative clauses in that they are not dominated by an NP headed by a null N. The case assignment domain remains clausal, yielding nominative to be assigned to the subject.

(18)   Adjunct clause:
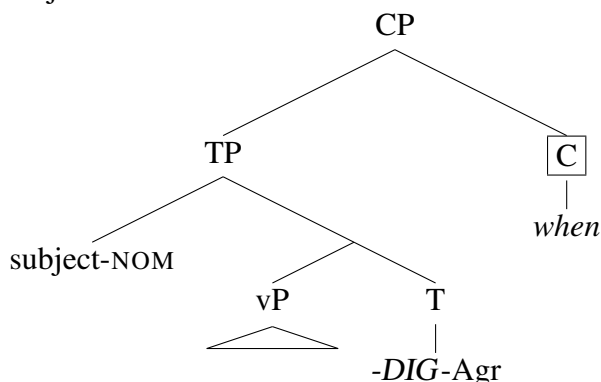
Let us now turn to some predictions of this proposal.

3.   Predictions

If the claim that argument clauses are complex NPs, but adjunct clauses are not, is correct, we expect that for suitable syntactic tests, argument clauses and complex NPs will pattern together to the exclusion of adjunct clauses. The following table illustrates the logically possible patterns in grammaticality, where α and β are variables standing for grammaticality statuses, and $\alpha \neq \beta$.

(19)   Possible patterns of tests

	Argument clause	Complex NP	Adjunct clause	Predicted
1	α	α	α	Yes
2	α	α	β	Yes
3	α	β	α	No
4	α	β	β	No

Pattern 1 is non-conclusive because it is neither evidence for the different structures in argument clauses and adjunct clauses, nor is it evidence against this hypothesis. Pattern 2 is predicted: argument clauses and complex NPs pattern together because argument clauses *are* complex NPs. Adjunct clauses pattern differently because they are not complex NPs. Patterns 3 and 4 are evidence against the present proposal (unless we can show that they are due to independent factors and actually instantiate pattern 1 or 2).

In this section I will run through some tests and show that none of them produce patterns 3 or 4. Two tests give us pattern 2, while the other two tests result in the inconclusive pattern 1.

(20)   Outcome of tests:

	Arg.cl.	CNP	Adj.cl.	Predicted	Attested
1	α	α	α	Yes	§3.1, §3.3
2	α	α	β	Yes	§3.2, §3.4
3	α	β	α	No	?
4	α	β	β	No	?

3.1. Topicalization fronting

Complex NPs have been reported to be islands for topicalization (Ross 1967). However, evidence from extraction patterns of topicalization is not conclusive: for topicalization of arguments, all three clause types allow extraction, and for topicalization of adverbs, neither clause type allows them. Complex NPs are thus not islands for scrambling in Turkish, and neither are argument clauses or adjunct clauses. The contrast between the extraction of arguments and of adverbs must be due to a different factor.

The following example illustrates the topicalization of the direct object *pilav-ı* 'rice-ACC' out of an argument clause (21a), a complex NP (21b), and an adjunct clause (21c). All extractions are possible.

(21)   a.   Pilav-ı Mehmet [Nurhan-ın   *t* ye-diğ-in   ]-i   düşün-üyor-du.
            rice-ACC Mehmet Nurhan-GEN   eat-*DIG*-3SG -ACC thnk-PROG-PST
            'Mehmet thought that Nurhan had eaten the rice.'
   b.   Pilav-ı Mehmet [[Nurhan-ın *t* ye-diğ-i   ] gerçeğ-in]-i
            rice-ACC M   N-GEN   eat-*DIG*-3SG fact-3SG -ACC
            bil-iyor-du.
            know-PROG-PST
            'Mehmet knew the fact that Nurhan had eaten the rice.'
   c.   Pilav-ı Mehmet [Nurhan-∅ *t* ye-diğ-i   için   ] üzül-dü.
            rice-ACC M   N-NOM   eat-*DIG*-3SG because sadden-PST
            'Mehmet was sad because Nurhan had eaten the rice.'

In contrast, topicalization of the adverb *dün* 'yesterday' is not allowed in argument clauses (22a), complex NPs (22b), and adjunct clauses (22c).

(22)   a.   *Dün   Mehmet [Nurhan-ın *t* pilav-ı   ye-diğ-in   ]-i   düşün-üyor.
            yesterday M   N-GEN   rice-ACC eat-*DIG*-3SG -ACC think-PROG
            'Mehmet thinks that Nurhan ate the rice yesterday.'
   b.   *Dün   M   [[Nurhan-ın *t* pilav-ı   ye-diğ-i   ] söylenti-ler-in]-e
            yesterday Mehmet N-GEN   rice-ACC eat-*DIG*-3SG   rumor-PL-3SG -DAT
            inan-ıyor.
            believe-PROG
            'Mehmet believes the rumors that Nurhan ate the rice yesterday.'
   c.   *Dün   Mehmet [Nurhan-∅ *t* pilav-ı   ye-diğ-i   için   ] bügün
            yesterday M   N-NOM   rice-ACC eat-*DIG*-3SG because   today
            açlık çek-iyor.
            go hungry-PROG
            'Because Nurhan ate the rice yesterday, Mehmet has to go hungry today.'

We can thus conclude that complex NPs in Turkish are not islands, so topicalization patterns cannot help us distinguish whether the hypothesis that argument clauses are complex NPs is correct.

3.2.   *Postverbal movement*

However, Turkish does not only have leftward scrambling, but also rightward movement to the postverbal position.[2] Postverbal phrases are usually associated with very weak prosody, and a 'givenness' or 'backgrounded' interpretation. For postverbal movement, we can observe that argument clauses (23a) and complex NPs (23b) allow this extraction, while adjunct clauses (23c) do not.[3]

(23) a. ?[Hasan-ın   nihayet *t* kaç-tığ-ın   ]-ı   duy-du-m   karı-sın-dan.
          Hasan-GEN finally   escape-*DIG*-3SG -ACC hear-PST-1SG wife-3SG-ABL
          'I heard that Hasan finally ran away from his wife.'

      b. ?[[Hasan-ın   nihayet *t* kaç-tığ-ı   ] söylenti-sin]-i   duy-du-m
          Hasan-GEN finally   escape-*DIG*-3SG   rumor-3SG -ACC hear-PST-1SG
          karı-sın-dan.
          wife-3SG-ABL
          'I heard the rumor that Hasan finally ran away from his wife.'

      c. *[[Hasan-ın   nihayet *t* kaç-tığ-ı   ] için   ] eğlen-di-k
          Hasan-GEN finally   escape-*DIG*-3SG   because celebrate-PST-1PL
          karı-sın-dan.
          wife-3SG-ABL
          'We celebrated because Hasan finally ran away from his wife.'

Argument clauses pattern thus with complex NPs, supporting the claim that they share the same structure in the relevant respect. Adjunct clauses, while not being islands for leftward topicalization of arguments in (22c), seem to not allow rightward movement of arguments (23c), indicating that they are not complex NPs. I leave further details of rightward movement for future research.

3.3.   *Wh-movement*

We can run three tests on wh-movement (relative clauses §3.3.1, pseudo-clefts §3.3.2, and sluicing §3.3.3), all of which produce the inconclusive pattern 1.

3.3.1.   *Relative clauses*

Relativization is allowed out of argument clauses (24a), out of complex NPs (24b), and out of adjunct clauses (24c).

---

[2] See Kural (1997), Kornfilt (2005), Şener (2010), Bošković & Şener (2014) for arguments based on c-command relations that postponing involves rightward movement, not base-generation or remnant movement.

[3] The judgments reported here are those of my consultants. Kornfilt (2003) reports (23b) as '??/*', which seems to indicate speaker variation. My consultants reported no difference in acceptability between (23a) and (23b), but they reported that (23c) was significantly more degraded than (23a,b).

(24)   a.   [Deniz-in   [biz-im   *t* al-dığ-ımız   ]-ı   san-dığ-ı   ] kitab-ı
                 Deniz-GEN   we-GEN   buy-*DIG*-1PL -ACC think-*DIG*-3SG   book-ACC
                 oku-du-m.
                 read-PST-1SG
                 'I read the book that Deniz thought we had bought.'

         b.   [Ali-nin   [[*t* oku-duğ-um   ] söylentiler-in]-i   duy-duğ-u   ] kitab-ı
                 Ali-GEN   read-*DIG*-1SG rumors-3SG   -ACC hear-*DIG*-1SG   book-ACC
                 al-dı-m.
                 read-PST-1SG
                 'I bought the book that Ali had heard the rumors that I had read.'

         c.   [Deniz-in   [biz-∅   *t* al-dığ-ımız   için   ] üzül-düğ-ü   ] kitab-ı
                 Deniz-GEN   we-NOM   buy-*DIG*-1PL because   sadden-*DIG*-3SG   book-ACC
                 oku-du-m.
                 read-PST-1SG
                 'I read the book that Deniz was sad because we had bought it.'

3.3.2.   Pseudo-clefts

Turkish pseudo-clefts consist of a free relative and a predicate with an often null copula (Kornfilt 1997b:192–3). Against the judgments reported in İnce (2005), my consultants allowed extraction out of complex NPs (25b) and adjunct clauses (25c), next to the uncontroversial extraction out of argument clauses (25a), yielding no difference between these clause types.

(25)   a.   [Ali-nin   [*t* oku-duğ-um   ]-u   duy-duğ-u   ] kitap.
                 Ali-GEN   read-*DIG*-1SG -ACC hear-*DIG*-1SG   book
                 'It's a book that Ali heard that I had read.'

         b.   [Ali-nin   [[*t* oku-duğ-um   ] söylentiler-in]-i   duy-duğ-u   ] kitap
                 Ali-GEN   read-*DIG*-1SG rumors-3SG   -ACC hear-*DIG*-1SG   book
                 'It's a book that Ali heard the rumors that I had read.'

         c.   [Ali-nin   [ben *t* oku-duğ-um   için   ] üzül-düğ-ü   ] kitap.
                 Ali-GEN   I   read-*DIG*-1SG because   sadden-*DIG*-3SG   book
                 'It's a book that Ali got sad because I read.'

3.3.3.   Sluicing

İnce (2005) argues that the particle *ki*, which expresses a meaning similar to 'I wonder …', attaches only to matrix questions. She observes that it can appear in sluices, which implies that there must be covert movement of the wh-operator to the matrix question in order to license *ki*. We can use this covert movement as a test for islandhood. My consultants reported that extraction is allowed out of all three clause types (contra the judgments reported in İnce (2005), where it was only possible for argument clauses).

(26)   a.   Hasan [Ahmet-in   birin-i   ara-dığ-ın   ]-ı   söyle-di.
                 Hasan   Ahmet-GEN one-ACC call-*DIG*-3SG -ACC say-PST-3SG
                 Kim-i   ki?
                 who-ACC *ki*
                 'Hasan said that Ahmet called someone. — Who, I wonder?' (İnce 2005:(16a))

    b. [[Ahmet-in   birin-i   vur-duğ-u   ] söylentiler-in]-i   duy-du-m.
         Ahmet-GEN one-ACC shoot-*DIG*-3SG   rumors-3SG -ACC hear-PST-1SG

         Kim-i   ki?
         who-ACC *ki*

         'I heard rumors that Ahmet had shot someone. — Who, I wonder?'

    c. Hasan [Ahmet birin-i   ara-dığ-ı   için   ] üzül-dü.
         Hasan   Ahmet one-ACC call-*DIG*-3SG because   sadden-PST

         Kim-i   ki?
         who-ACC *ki*

         'Hasan became sad because Ahmet called someone. — Who, I wonder?'

In this section, we have seen three types of wh-movement, all of which seemed to not be sensitive to clause type. Argument clauses, complex NPs, and adjunct clauses all allow extraction of wh-elements.

3.4.   Head noun insertion test

As Aygen (2007) observes, the proposal that argument clauses are complex NPs while adjunct clauses are not, makes another clear prediction: the covert head noun should be able to be optionally overt. Indeed, the noun can be made overt in argument clauses, turning (27a) into the complex NP in (27b).

(27)   a. Ben [[Hasan-ın   bu kitab-ı   oku-duğ-u   ] ∅]-nu bil-iyor-um.
          1SG   Hasan-GEN this book-ACC read-*DIG*-3SG   -ACC know-PROG-1SG
          'I know that Hasan read this book.'

      b. Ben [[Hasan-ın   bu kitab-ı   oku-duğ-u   ] gerçeğ-in]-i
          1SG   Hasan-GEN this book-ACC read-*DIG*-3SG   fact-3SG   -ACC

          bil-iyor-um.
          know-PROG-1SG
          'I know the fact that Hasan read this book.'

Note that in (27a), the silent head noun's accusative marker is pronounced, while the compound marker 3SG disappears.

    Inserting a head noun is not possible for adjunct clauses, indicating that they are not dominated by an NP. (28a) is a grammatical adjunct clause with nominative on the subject. In (28b), a head noun is inserted, making the sentence ungrammatical. Following Aygen (2007), I take *için* to be homophonous between a complementizer and a postposition. That neither option is available here is shown in (28b) and (28c).

(28)   a. Ben [Hasan-∅   bu kitab-ı   oku-duğ-u   için   ] üzül-dü-m.
          1SG   Hasan-NOM this book-ACC read-*DIG*-3SG because   sadden-PST-1SG
          'I became sad because Hasan read this book.'

      b. *Ben [[Hasan-∅   bu kitab-ı   oku-duğ-u   için   ] gerçeğ-i]
          1SG   Hasan-NOM this book-ACC read-*DIG*-3SG because   fact-3SG

          üzül-dü-m.
          sadden-PST-1SG

intended: 'I became sad because of the fact that Hasan read this book.'

c. *Ben [[Hasan-∅   bu kitab-ı   oku-duğ-u   ] gerçeğ-i] için
    1SG   Hasan-NOM this book-ACC read-*DIG*-3SG   fact-3SG   because
    üzül-dü-m.
    sadden-PST-1SG
    intended: 'I became sad because of the fact that Hasan read this book.'

As a minimal contrast to (28c), (29a) is grammatical, but the subject is marked with genitive, as we expect given that it is a complex NP.

(29)   Ben [[Hasan-ın   bu kitab-ı   oku-duğ-u   ] gerçeğ-i] için
        1SG   Hasan-GEN this book-ACC read-*DIG*-3SG   fact-3SG   because
        üzül-dü-m.
        sadden-PST-1SG
        'I became sad because of the fact that Hasan read this book.'

The same effects of homophonous C/P alternations is illustrated in the following examples from Aygen (2007). The postposition *göre* has the meaning 'according to' and selects NPs, assigning lexical dative case:

(30)   Hasan-a   göre
        Hasan-DAT *göre*
        'according to Hasan'

In (31) the postposition *göre* 'according to' selects a free relative with -*DIG*. We can confirm that this is a relative clause by testing for the gap (31b) and by making the head noun overt (31c). As expected, the subject carries genitive case.

(31)   a. [Hasan-ın   duy-duğ-un   ]-a   göre herkes   duy-acak-mış.
            Hasan-GEN hear-*DIG*-3SG -DAT *göre* everyone hear-FUT-REP
            'According to what Hasan heard, everybody will hear it.'
        b. [Hasan-ın   (*on-u) duy-duğ-un   ]-a   göre herkes duy-acak-mış.
            Hasan-GEN   it-ACC hear-*DIG*-3SG -DAT *göre*
            'According to what Hasan heard (*it), everybody will hear it.'
        c. [[Hasan-ın   *t* duy-duğ-u   ] şey   ]-e   göre herkes duy-acak-mış.
            Hasan-GEN   hear-*DIG*-3SG thing -DAT *göre*
            'According to the thing that Hasan heard, everybody will hear it.'

In (32a), the complementizer *göre* translates as 'since'/'because'. The embedded subject carries nominative case. Under this interpretation, it does not select for an NP, as we can test for in (32b), where there is no gap as in the relative clauses above, and in (32c–d), where insertion of a head noun is not possible.

(32)   a. [Hasan-∅   duy-duğ-un-a   göre] herkes   duy-acak.
            Hasan-NOM hear-*DIG*-3SG-DAT *göre*   everyone hear-FUT
            'Given that Hasan heard it, everyone will hear it.'
        b. [Hasan-∅   (haber-i) duy-duğ-un-a   göre] herkes duy-acak.
            Hasan-NOM news-ACC hear-*DIG*-3SG-DAT *göre*
            'Given that Hasan heard the news, everyone will hear them.'

c. *[[Hasan-∅   duy-duğ-u   ]-na   göre şey   ] herkes duy-acak.
     Hasan-NOM hear-*DIG*-3SG   -DAT *göre* thing

d. *[[Hasan-∅   duy-duğ-u   ] şey   -e   göre] herkes duy-acak.
     Hasan-NOM hear-*DIG*-3SG   thing -DAT *göre*

Again, (32d) would be a grammatical headed relative clause if the subject were marked with genitive, cf. (31c).

4. Conclusions

In this paper, I have presented an analysis that extends a configurational case system to clause-type based DSM in Turkish. Building on Lees's (1965) and Aygen's (2007) observation that argument clauses are headed by a silent abstract noun, I have shown that this analysis accounts for the observed DSM-patterns when we pair it with a configurational theory of case assignment. Together with the evidence for requiring a configurational system in the first place (see section 2.1), this makes it unnecessary to argue for a heterogeneous system where a single language employs both strategies (Baker & Vinokurova 2010, Gribanova 2016; see also Levin & Preminger 2015 for counter-arguments), and also goes against pure Chomsky (2000, 2001)-style approaches of case licensing (Kornfilt 2001, 2003, 2006; i.a.).

I have laid out some predictions that this account makes and shown that there are two syntactic tests that fulfill them (postverbal movement §3.2, head noun insertion §3.4). Two other tests that have been proposed for parallel structures in other languages (topicalization §3.1, wh-movement §3.3) are inconclusive in Turkish. At this point, I am not aware of any tests that falsify my proposal.

The present theory contrasts with Colley & Davis's (to appear) proposal, who claim that adjunct clauses involve more projections than argument clauses. However, the head noun insertion test (§3.4) shows that the silent N cannot be made overt in adjunct clauses, thereby suggesting that it is not present.

Acknowledgements

Sabine Laszakovits
University of Connecticut
sabine.laszakovits@uconn.edu
https://sabine.laszakovits.net

References

Aydemir, Y. (2004). Are Turkish preverbal bare nouns syntactic arguments? *Linguistic Inquiry* 35, pp. 465–474.

Aygen, G. (1999). Specificity and Subject-Object Positions: Scope Interactions in Turkish. ms, Harvard University.

Aygen, G. (2007). Syntax and Semantics of Genitive Subject-Case in Turkish. *California Linguistic Notes* 32:2.

Baker, M. (2015). *Case: Its principles and its parameters*. Cambridge University Press.

Baker, M. & N. Vinokurova (2010). Two modalities of case assignment: case in Sakha. *Natural Language & Linguistic Theory* 28:3, pp. 593–642.

Bittner, M. & K. Hale (1996). The structural determination of Case and Agreement. *Linguistic Inquiry* 27, pp. 1–68.

Bossong, G. (1985). *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Gunter Narr Verlag, Tübingen.

Bošković, Ž. & S. Şener (2014). The Turkish NP. Hofherr, P. C. & A. Zribi-Hertz (eds.), *Crosslinguistic studies on nominal reference: With and Without Articles*, Brill, Leiden, pp. 102–140.

Chomsky, N. (2000). Minimalist inquiries: the framework. Martin, R., D. Michaels & J. Uriagereka (eds.), *Step by step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, MIT Press, pp. 89–155.

Chomsky, N. (2001). Derivation by phase. Kenstowicz, M. (ed.), *Ken Hale: a life in language*, MIT Press, pp. 1–52.

Colley, J. & C. Davis (to appear). A new approach to Turkish nominalized clauses. Zidani-Eroğlu, L., M. Ciscel & H. Koulidobrova (eds.), *Proceedings of WAFL 12*.

den Dikken, M. (2007). Phase Extension Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33, pp. 1–41.

Enç, M. (1991). The Semantics of Specificity. *Linguistic Inquiry* 22:1, pp. 1–25.

Gallego, Á. (2010). *Phase theory*. John Benjamins.

Göksel, A. & C. Kerslake (2005). *Turkish. A Comprehensive Grammar*. Routledge.

Gribanova, V. (2016). Case, agreement, and differential subject marking in Uzbek nominalized clauses. Ms. Stanford.

von Heusinger, K. & J. Kornfilt (2005). The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages* 9, pp. 3–44.

İnce, A. (2005). Island-sensitive sluicing in Turkish. Bateman, L. & C. Ussery (eds.), *Proceedings of NELS 35*, pp. 279–290.

Kornfilt, J. (1997a). On the Syntax and Morphology of Relative Clauses in Turkish. İmer, K., A. Kocaman & S. Özsoy (eds.), *Dilbilim Araştırmaları (Linguistic Investigations)*, Kebikeç Yayınları, Ankara, pp. 24–51.

Kornfilt, J. (1997b). *Turkish*. Descriptive Grammars, Routledge, London & New York.

Kornfilt, J. (2001). Functional projections and their subjects in Turkish clauses. Erguvanlı Taylan, E. (ed.), *The Verb in Turkish*, John Benjamins, vol. 44 of *Linguistik Aktuell*, p. 183–212.

Kornfilt, J. (2003). Subject Case in Turkish Nominalized Clauses. Junghanns, U. & L. Szucsich (eds.), *Syntactic Structures and Morphological Information*, Mouton de Gruyter, pp. 129–215.

Kornfilt, J. (2005). Asymmetries between pre-verbal and post-verbal scrambling in Turkish. Sabel, J. & M. Saito (eds.), *The free word order phenomenon: its syntactic sources and diversity*, Mouton de Gruyter, Berlin, pp. 163–181.

Kornfilt, J. (2006). Agreement: The (unique and local) syntactic and morphological licenser of subject case. Costa, J. & M. C. Figueiredo (eds.), *Studies on agreement*, John Benjamins, pp. 141–171.

Kornfilt, J. & O. Preminger (2015). Nominative as 'no case at all': an argument from raising-to-ACC in Sakha. Joseph, A. & E. Predolac (eds.), *Proceedings of the 9th Workshop on Altaic Formal Linguistics (WAFL9)*, MIT Press, vol. 76 of *MIT Working Papers in Linguistics*, pp. 109–120.

Kural, M. (1997). Postverbal constituents in Turkish and the Linear Correspondence Axiom. *Linguistic Inquiry* 28, pp. 498–519.

Lees, R. (1965). Turkish Nominalizations and a Problem of Ellipsis. *Foundations of Language* 1:2, pp. 112–121.

Levin, T. & O. Preminger (2015). Case in Sakha: are two modalities really necessary? *Natural Language & Linguistic Theory* 33:1, pp. 231–250.

Marantz, A. (1992). Case and licensing. Westphal, G., B. Ao & H.-R. Chae (eds.), *Proceedings of the Eastern States Conference on Linguistics (ESCOL) 1991*, Cornell University, CLC Publications, Ithaca, NY, pp. 234–253.

Nilsson, B. (1986). Object Incorporation in Turkish. Koç, A. & E. Erguvanlı Taylan (eds.), *Türk Dilbilimi Konferansı Bildirleri, 9–10 Ağustos 1984. Proceedings of the Turkish Linguistics Conference, August 9–10, 1984*, Boğaziçi University Publications, Istanbul, pp. 113–128.

Ross, H. (1967). *Constraints on Variables in Syntax*. Ph.D. thesis, MIT.

Şener, S. (2008). Non-Canonical Case Licensing is Canonical: Accusative Subjects of CPs in Turkish. URL http://ling.auf.net/lingbuzz/000759. Ms, UConn.

Şener, S. (2010). *(Non-)Peripheral Matters in Turkish Syntax*. Ph.D. thesis, University of Connecticut.

Şener, S. (2011). Cross Clausal Licensing of Accusative Case on Subjects of CPs in Turkish. Lima, S., K. Mullin & B. Smith (eds.), *NELS 39: Proceedings of the 39th Annual Meeting of the North East Linguistic Society*, CreateSpace Independent Publishing Platform, pp. 679–690.

Wurmbrand, S. (2013). QR and selection: Covert evidence for phasehood. Keine, S. & S. Sloggett (eds.), *Proceedings of the North Eastern Linguistics Society Annual Meeting 42 (NELS 42)*, University of Massachusetts, GLSA, Amherst, pp. 619–632.

Yip, M., J. Maling & R. Jackendoff (1987). Case in tiers. *Language* 63, pp. 217–250.

Parametrizing intralinguistic variation:
Case assignment strategies in Russian event nominalizations[1]

Anastasia Gerasimova

The paper deals with the issue of intralingual variation in the context of Russian event nominalizations. Experimental data show that external arguments in nominalizations can be assigned both genitive and instrumental case, i.e. exhibit differential case marking. In this paper I investigate the distribution of the two patterns of case assignment among speakers and demonstrate that both configurations can coexist in one individual grammar. My analysis suggests that it is the amount of functional structure of the constituent undergoing nominalization that determines the observed case variation.

1. Introduction

The concept of parametric variation is foundational in linguistics. Since the Theory of Principles and Parameters it has been common to see the differences between languages as the result of parameter realization (Chomsky & Lasnik 1993; Pesetsky 2003). This approach was preserved within the Minimalist framework, where formal features of lexical and functional heads serve as parameters predetermining the outcome of the derivation. Consequently, cross-linguistic morphosyntactic variation results from language-specific values of general parameters.

However, variation can also be found within a single language. Broadly, one can distinguish between two types of intralingual variation; either grammatical variants of one language are distributed among the speakers, or they coexist within individual grammars. In both cases the variation can be parameterized: either the parameter settings are different for distinct speech communities (e.g. as in the case of dialects or sociolects) or the parameter settings may be triggered by factors internal to the grammar.

This paper addresses the issue of intralingual variation in the context of Russian event nominalizations. Nominalizations are verbal derivatives that demonstrate both nominal and verbal properties. Abney (1987) proposed that the syntactic structure of nominalizations includes a verbal phrase, which explains the derivational relation with verbs and such verbal properties of nominalizations as case assignment and adverbial modification. Since the

external distribution of nominalizations is identical to the distribution of noun phrases, the verbal phrase is generally analyzed as being embedded within a DP projection.

Grimshaw (1990) provided evidence for the existence of two types of nominalizations, namely result and process (event) nominals. Both types have an argument structure which is associated with the VP. The distinction can be proven by a set of diagnostics. For example, process nominals obligatorily take internal arguments (1a), and the telicity test is also applicable to them. Only nominals with a process interpretation may take aspectual modifiers (consider (1b) vs. (1c)) and agent-oriented modifiers (1d),

(1)   a. Vashe   upravlenie   *(**pomest'-em**)   prodlitsja
           your   management (PROCESS)   estate-INSTR   will last
           do vstuplenija   naslednika   v   svoi   prava.
           till coming   heir.GEN   into   his   rights
           'Your management of the estate will continue till accession of the heir.'

        b. **Postojannoe**   otkryvanie   shkaf-a   privelo
           constant   opening (PROCESS)   wardrobe-GEN   lead
           k tomu,   chto   ruchka   otvalilas'
           to   that   handle   fell off
           'The constant opening of the wardrobe resulted in that the handle fell off.'

        c. (*postojannoe*)   otkrytie   schet-a v banke   pozvoljaet
           constant   opening (RESULT)   bank account-GEN   allows
           vam   uchastvovat'   v   rozygryshe
            you   participate   in   give-away
           'The constant opening of the bank account allows you to participate in a give-away.'

        d. nanesenie   im   Tarasovu
           causing (PROCESS)   he.INSTR   Tarasov.DAT
           umyshlenno   telesn-yh   povrezhdeni-j
           deliberately (AGENT-ORIENTED MODIFIER)   bodily-GEN   injuries-GEN
           'his causing injury to Tarasov in a deliberate way'
           (Pazel'skaya & Tatevosov 2008: (8))

Alexiadou (2001) proposed that the differences between the two types of nominalizations result from differences in the amount of functional structure dominating the VP: process nominals, unlike result nominalizations, contain at least the *v*P projection. Alexiadou shows that higher verbal projections can also be present in the syntactic structure of nominalizations: thus, manner modification correlates with the presence of the aspectual layer AspP in Polish; while the embedding of a CP results in a nominalized clause in Greek. Therefore, the structural shape of nominalizations is a parameter of cross-linguistic variation.

The amount of functional structure that is nominalized can also be a parameter of intralinguistic variation. In this paper I focus on case marking patterns within Russian nominalizations and investigate whether the variation found within Russian nominalizations is the result of different parameter settings. The aim of this study is to show what triggers the

variation: the case marking strategy that speakers stick to, or the functional "size" of the nominalization.

The structure of the paper is as follows. In Section 2, I provide an overview of Russian event nominalizations and consider case marking strategies with different structural types of verbal stems. Section 3 presents an experimental study on what strategy speakers choose in the process of speech production and defines the range of the case variation. Next, using experimentally obtained acceptability judgments I study the availability of the two case marking strategies for different speakers (Section 4). Finally, Section 5 provides an analysis of differential case marking in Russian event nominalizations: I suggest that case variation follows from the variable structural shape of Russian event nominalizations.

2. Argument structure of Russian event nominalizations

In Russian, eventive nominalizations include regular deverbal nominals derived with the highly productive suffixes *-nij/-tij-*, as well as nominals formed with less productive affixes *-k-*, *-ot-*, *-stv-*, *-b-* and *-zn'-* (Shvedova 1980). In this paper I adopt the approach proposed by Pazel'skaya and Tatevosov (2008), who argue that *-nij-/-tij-* are in fact complex affixes that consist of the "substantive" affix *-ij-*, and the "nominal" affix *-n-/-t-*, which can be also found within passive participles. I will also rely on the Mirror Principle that states that affix ordering reflects the sequence of syntactic derivations (Baker 1985).

As noted above, process nominalizations have an argument structure which is associated with functional layers. The *v*P projection is present within process nominals only. Pazel'skaya and Tatevosov (2008) argue that the presence of the *v*P layer in Russian process nominalizations correlates with several syntactic properties, viz.: availability for adverbial modification (e.g. agent-oriented modifiers), causative-inchoative alternation, and the ability of the external argument to control PRO in purpose clauses (Roeper 1987).

As for the higher verbal projections, Alexiadou (2001) compares Russian and Polish nominals and states that there is no AspP in Russian nominalizations because there is no true aspectual opposition. Pazel'skaya and Tatevosov (2008) investigate the set of affixes that are possible within the nominalized stem and their compatibility. Contrary to Alexiadou, they show that nominalizations may attach the imperfective suffix *-yva-*, which indicates the presence of AspP. Therefore, the maximum projection which can appear in Russian process nominalizations is AspP.

Lyutikova (2014) develops the hypothesis that nominalizations have argument structure regardless of the presence of derivational affixes in the stem. In other words, nominals which lack *-nij-/-tij-* or other suffixes are still considered to be derived from the verb stem (e.g. *torgovlja* 'trade', *kritika* 'criticising'). In this paper I will follow this view on the structure of nominalizations.

2.1. Case marking strategies

The case marking strategy for arguments of nominalizations is a parameter of cross-linguistic variation in WALS (Koptjevskaja-Tamm 2013). The classification created by Koptjevskaja-Tamm (2002) is based upon differences in marking of the only argument of an intransitive stem (S), agent (A) and patient/theme (P) of a transitive stem. Russian event nominalizations

belong to the ergative-possessive type. This means that the arguments of intransitives and internal arguments of transitive stems are marked with the possessive, genitive case (GEN), while external arguments of transitives are assigned instrumental case (INSTR).

(2)   a. Internal argument of intransitive marked GEN
          padenie   kurs-a   rubl-ja
          fall   course-GEN   rouble-GEN
          'fall of rouble course (weakening of rouble)'

      b. External argument of intransitive marked GEN
          vorchanie   sosed-ej
          grumbling neighbors-GEN
          'neighbors' grumbling'

      c. Internal argument marked GEN and external argument marked INSTR within transitive stem
          ispolnenie   ari-i   Shaljapin-ym
          performance   aria-GEN   Chaliapin-INSTR
          'performance of aria by Chaliapin'

Nevertheless, the corpus and Internet data show that instrumental case marking of the external arguments is not limited to prototypical transitive stems (Pereltsvaig, Lyutikova, Gerasimova 2016, forthcoming). Firstly, the external argument of transitive nominalizations with lexically governed internal arguments can be marked with both the GEN and INSTR (3-4).

(3)   podrazhanie   chelovek-a   prirod-e
      imitating   man-GEN   nature-DAT
      'man's imitation of nature'

(4)   podrazhanie   chelovek-om tvorchesk-oj moshch-i
      imitating   man-INSTR   creative-DAT power-DAT
      'man's imitation of the creative power'

Secondly, the INSTR is also possible with intransitive unergative stems, stems that do not have internal argument.

(5)   hozhdenie   rebenk-om   na cypochkah
      walking.around   child-INSTR   on tiptoe
      'child's walking on tiptoe'

Two alternatives are possible, which means that the external argument demonstrates differential case marking. In other words, the case marking strategy is one of the parameters of intralingual variation for Russian. However, corpus data provide no information either on speakers' consistency in assigning INSTR to external arguments or on the quantitative characteristics of INSTR usage with different stems. Therefore, it is necessary to conduct an experiment that would provide evidence for the existence of two case marking strategies in the modern Russian language.

3. Differential case marking

In order to define the range of variation, I conducted an experiment focusing on what cases speakers choose in the process of speech production. The examined stems included nominalizations derived from unergative stems and transitive stems with lexical government. The latter include nominals with internal argument in dative (Trans-DAT), instrumental (Trans-INSTR), or in the form of prepositional phrase (Trans-PP). The nominalizations with an internal argument in the GEN were examined in the experiments, however, they were not considered in the analysis since speakers tend to avoid the same case on both internal and external arguments. In other words, for this type of nominal the external argument is assigned INSTR due to other factors than for other types.

To verify the results of the experiment the stimuli also included unaccusative nominalizations, which have only internal arguments, and prototypical transitive stems.

The data was collected from 120 participants. The study was conducted via Google Forms and consisted of a fill-in-the-blanks task: native speakers were asked to generate arguments of nominalizations assigning cases that sounded most natural to them. Examples of stimuli for different stems can be seen in (6) (verbs and nominalizations are italicized). An example of the filled gap can be found in (7).

(6)   a. Unaccusative stem
V jetom godu **prezident** *pribyl* na final'nyj match, i *pribytie* _____ privleklo mnogo vnimanija k chempionatu.
'This year the president.**NOM** *arrived* to the final game, and the *arrival* _____ drew attention to the championship.'

b. Unergative stem
Dazhe zimoj **dedushka** *plavaet* v rechke kazhdyj den', hotja ezhednevnoe *plavanie*_____ bespokoit vsju sem'ju.
'Even in the winter the **grandfather.NOM** *swims* in the river every day, although day-to-day *swimming* _____ worries the whole family.'

c. Trans-DAT stem
V ljubom spore **sekretar'** terpelivo *poddakival* **nachal'nik-u**, no ezhednevnoe poddakivanie _____ _____ ni k chemu ne privelo.'
'In any argument the **secretary.NOM** *has* patiently *consented* to the director-**DAT**, but the everyday *consenting* _____ _____ lead to nowhere.'

d. Trans-INSTR stem
Vse leto **sosedka** *torgovala* **kartin-ami** v centre goroda, no torgovlja _____ _____ prinosila mizernyj dohod.
'The whole summer the **neighbour.NOM** *traded* paintings-**INSTR** in the city center, but the *trade* _____ _____ was bringing a poor income.'

e. Trans-PP stem
Celyj god **inzhener** terpelivo *uhazhival* za **artistk-oj**, hotja uhazhivanie _____ _____ bylo ves'ma svoeobraznym.

'The whole year the **engineer.NOM** *has* patiently *been courting* the **actress-INSTR (PP)**, but the *courtship* _____ _____ was pretty strange.'
f. Transitive stem
V tot mesjac **armija** *osvobodila* **stolic-u**, i *osvobozhdenie* _____ _____ sil'no podnjalo boevoj duh vseh soldat.
'That month the **army.NOM** *reconquered* the **capital-ACC**, and *reconquest* ____ ____ lift the martial spirit.'

(7)   V tot mesjac **armija** osvobodila **stolic-u**, i osvobozhdenie *armiej stolicy* sil'no podnjalo boevoj duh vseh soldat.
'That month the **army.NOM** reconquered the **capital-ACC**, and reconquest by the army.INSTR of the capital.GEN lift the martial spirit.'

8)   В тот месяц **армия** *освободила* **столицу**, и *освобождение* _____ _____ сильно подняло боевой дух всех солдат.

8) *

[   ]

Figure 1. Example fill-in-the-blanks task for the sentence (6f) as it was in the Google Form

The results of the experiment validate the hypothesis about two case marking strategies in Russian nominalizations. On the basis of the experimental data we can conclude that in this experiment Russian native speakers are inconsistent in using INSTR. In accordance with the observed usage speakers can be grouped into two clusters depending on how frequently they used INSTR in the survey.

Among the 120 participants 56% of respondents assigned INSTR to external arguments of nominalizations with lexically governed internal argument at least once. There was no significant correlation between age, gender, city, level of education and the ability to use INSTR. Consequently, case marking strategy is indeed one of the parameters of intralingual variation for Russian.

The results have also shown that within intransitive stems the external argument is commonly marked with GEN. In unergatives, external arguments are marked INSTR significantly rarely, but such cases exist (e.g. *plavanie dedushkoj* 'swimming of the grandfather.GEN'). Finally, for transitive nominalizations the ergative-posessive model was confirmed.
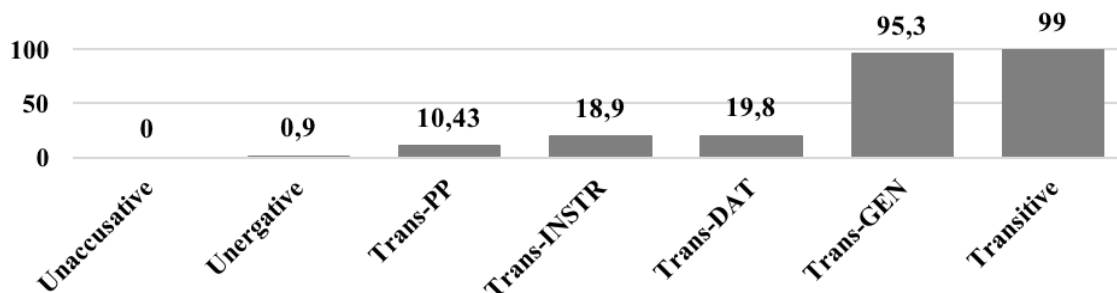
Figure 2. INSTR case marking (percentage %)

3.1. What case is predicted?

As was previously shown in Pereltsvaig, Lyutikova & Gerasimova (2016), the two existing case theories make distinct predictions about case marking for nominalizations whose internal argument is lexically governed.

Dependent Case Theory (Marantz 1991) predicts that the external argument is assigned GEN. In this theory case is considered to be a morphological phenomenon that is determined by the structural position of a caseless DP and other DPs within a case competition domain. Case realization is predicted by the disjunctive hierarchy, according to which at the beginning of the derivation most specific case features take precedence over everything else. In this theory Russian GEN is an unmarked adnominal case, which appears under the lack of another caseless DP within a case competition domain. The internal argument is assigned lexically-governed case. It is impossible to assign the dependent case because the internal argument is already marked. Consequently, following the case realization disjunctive hierarchy the external argument receives unmarked case.

On the contrary, Inherent Case Theory (Woolford 2006) expects the external argument to be marked with INSTR. This theory distinguishes two types of nonstructural Case: lexical Case and inherent Case, that are in complementary distribution. In Russian INSTR is an inherent Case, which is assigned independently from the internal argument and which is associated with the Agent θ-role.

For unergative stems we can find the same predictions. For unergatives the internal argument cannot influence case assignment mechanisms because it does not exist. In other words, in Dependent Case Theory there is no other caseless DP in the case competition domain. In Inherent Case Theory the assignment of the inherent Case does not correlate with the assignment of the lexical Case. Although the respondents have used external argument of unergative stems in INSTR significantly infrequently, from the predictions of the two theories and from earlier examples from the Internet (4) we can draw the conclusion that differential case marking may occur within unergative nominalizations.

So far we have examined two opposing theories, both of which can be used for modelling different phenomena. Each of the theories can justify only one alternative in terms of differential case marking. However, would it be legitimate to conclude that neither of the theories corresponds to the language reality? To answer this question we should take into account the different types of intralingual variation.

3.2. Types of intralingual variation

Variation commonly exists within a single language and falls into two types of variance: grammatical variants of one language can be distributed among the speakers or they may coexist within the grammatical scope of one individual. The first type of variation prevails within speech communities and is dictated by a number of factors such as area, social class, gender, and genre, such as the *was*/*were* variation across English dialects and sociolects (Anderwald 2001). Another type occurs in the speech of one individual and is caused by the presence of several variants for one linguistic configuration. In this case the choice of the grammatical variant is determined by a factor internal to the grammar.

In order to model the differential case marking of the external argument it is essential to find out which model of intralingual variation is represented in the case of Russian event nominalizations. If the ability to mark the external argument with INSTR is individual, then the two modalities of case assignment will be distributed among speakers. Alternatively, if both strategies are equally available to any speaker, we should account for how the two case configurations can coexist in one individual grammar. In other words, the case assignment mechanisms have to allow the choice between GEN and INSTR. The problem is particularly interesting in the context of the recent discussion as to whether the coexistence of two modalities of case assignment is possible within one language (Baker & Vinokurova 2010; Levin & Preminger 2015).

The experimental data has shown the inconsistency of Russian native speakers in using INSTR: more than 50% of respondents have assigned INSTR to the external argument of nominalizations with lexical government. To this end I suggest a hypothesis that the two modalities of case assignment in Russian nominalizations are distributed among the speakers.

4. Distribution of two modalities of case assignment

In order to test the hypothesis I conducted a new experiment with the same respondents. This experiment examined how speakers estimate the acceptability of cases when reading sentences. The data was collected from 78 respondents from the first experiment. The task involved evaluating the acceptability of event nominal constructions with external argument marked GEN or INSTR using a five point Likert scale. The examined stems included: transitives, transitive stems with lexically governed internal argument, and unergatives.

In the first experiment INSTR was rarely assigned to external arguments of unergative nominalizations. However, such examples have been found in the corpus. Interestingly, most of the examples found are characterized by an adverbial or PP modification of the nominalization. Hence, along with sentences with bare unergative nominalizations I added sentences with the same unergative stems modified with adverbs to the stimuli.

(8)   a. *Gracioznoe hozhdenie model-ej bylo vysoko oceneno dizajnerom.*
             'Graceful walking of the models-GEN was highly appreciated by the designer.'

         b. *Gracioznoe hozhdenie modeljami **po podiumu** bylo vysoko oceneno dizajnerom.*
             'Graceful walking by the models.INSTR **on the runway** was highly appreciated by the designer.'

Figure 3. Evaluation task for the sentence (8a) as it appeared in the Google Form

As the result of the experiment we analyzed the acceptability judgements of the stimuli and calculated the average acceptability scores.[2]

The experimental study revealed that there is a significant difference in INSTR acceptance for different stems.[3] Critically, INSTR is most acceptable with transitives. It is significantly less acceptable with stems with lexical government. Finally, INSTR was considered to be the least acceptable within unergative stems. This distribution of the scores correlates with the results of the first experiment, except for unergatives. Within the latter INSTR was practically not used, however, the acceptability score was significantly higher than the lowest possible score. Another remarkable result to emerge from the data is that PP modification of unergative stems significantly increases the acceptability of the external argument in INSTR. The difference between the scores of INSTR with modified unergatives and transitives with lexical government appeared to be non-significant. Consequently, it can be concluded that PP-modification raises the acceptability of INSTR.
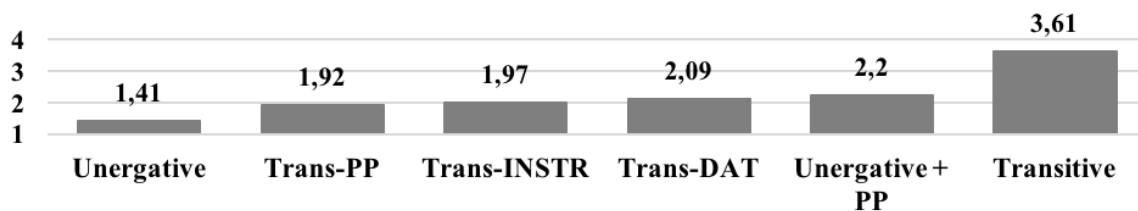
Figure 4. Acceptability of INSTR case marking (mean values)

Another intriguing correlation is related to the consistency of judgements of INSTR. Speakers can be grouped into two clusters depending on how they evaluated INSTR in the second experiment.[4] In other words, the survey has shown that Russian native speakers are inconsistent in using INSTR.
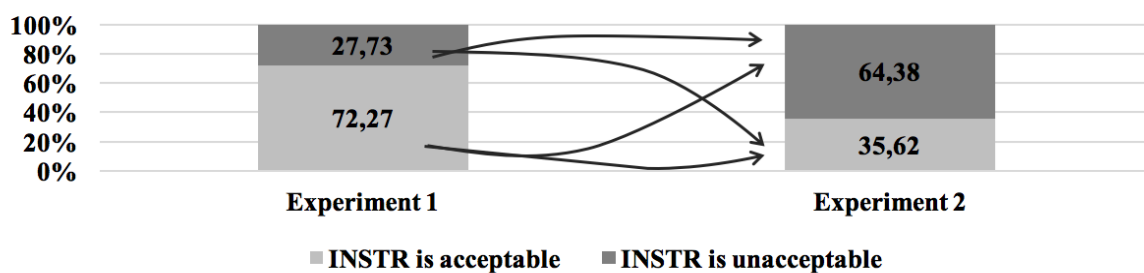
Figure 5. Respondents' grouping based on INSTR acceptability

---

[2] Prior to analyzing the average scores I examined the consistency of the range respondents' judgements. To do this we compared the variance of scores of each respondent with the group variance using the Fisher test. The test shows whether the ratio of the two variances goes beyond stochastic fluctuations. As a result I found seven respondents who demonstrated significantly low variance, and two respondents who demonstrated significantly high variance (had a lot of polar scores), which makes only 10% of all the participants. The other 90% demonstrated statistically the same range of scores.

[3] To group the types of nominalizations Student's *t*-test was applied to the sets of acceptability scores for different types of eventive nominalizations. *T*-test shows whether the mean differences between the sets are significant. In order to avoid an assumption about the normality of variance the sets of scores have also been tested for homogeneity using Mann-Whitney *U*-test. The results of the two statistical tests coincided.

[4] To find clusters, the *k*-means clustering method (*k*=2) was applied.

In the next stage I compared the acceptability judgements with data on production, namely whether the respondent assigned INSTR to the external argument of stems that allow for that. Curiously, speakers' evaluation of the INSTR acceptability is not consistent with the grouping based on actual usage in speech. Thus, the participant who constantly assigns INSTR to external arguments of action nominal constructions may estimate sentences with INSTR to be much less preferable than with GEN and vice versa. This means that the two strategies of case assignment to arguments of nominalization are equally available to any speaker. The hypothesis that predicts in Russian two modalities of case assignment distributed among the speakers must be rejected. Therefore, the experimental study shows that there is a need for a new model that would take into account an eventual coexistence of two alternatives for case marking and explain speakers' preferences when choosing the case of the external argument.

5. Analysis

It was previously said that the structural shape of nominalization is one of the parameters of cross-linguistic variation (Alexiadou 2001). This analysis will go a step further and build on the idea that the structure of nominalizations can vary within one language.

Kratzer (1996) claims that English nominalizers can embed syntactic structures of variable size. She compared the two types of gerunds (acc_{ing} and $poss_{ing}$) in which the direct object can receive accusative case and the type of gerund (of_{ing}) that attaches the direct object via the preposition *of*. Kratzer claims that the difference in argument structure results from the fact that acc_{ing} and $poss_{ing}$ gerunds embed vP, while of_{ing} contains at most VP.

(9)   a. his rebuilding of the barn (of_{ing})   (Kratzer 1996:33)
        b. his rebuilding the barn   ($poss_{ing}$)

The same idea was implemented in the analysis of the causative-inchoative alternation in Russian. Tatevosov (2008) argues that its interpretation depends on the amount of structure that is nominalized. If it is the VP that is nominalized, the semantic interpretation is intransitive, whilst when the vP is nominalized, the nominalization has transitive meaning.

I will use the same approach to model the differential case marking in Russian nominalizations. First, we will consider the two opposing options. An external argument in INSTR is obligatory with transitives but impossible with unaccusatives, which do not have an external argument that could be assigned INSTR. There is clear evidence for the vP in the structure of transitive nominalizations. Otherwise it would be impossible to express both internal and external arguments. On the contrary, with unaccusatives there is no external argument. Consequently, the functional head v has no purpose in appearing within unaccusatives.

Transitive stems with lexical government and unergatives are similar in terms of case assignment mechanisms. Then the differential case marking is determined by the amount of structure that is nominalized: the external argument is assigned INSTR when the nominalizer embeds the vP, and it is assigned GEN when the nominalizer embeds the VP.

The validity of the proposed model is emphasized by the syntactic behavior of unergatives. With unergative stems PP-modification increases the acceptability of INSTR. This argues for a greater amount of syntactic structure. Moreover, the structural differences between the two

types of nominalizations can be proved by the ungrammaticality of conjunction-reduction (10).

(10)   a. *poddakivanie   sekretar-ja   director-u   i
          consenting   secretary-GEN   director-DAT   and
          buhgalter-om   zamestitel-ju
          accountant-INSTR   deputy-DAT

       b. poddakivanie   sekretar-ja   director-u   i
          consenting   secretary-GEN   director-DAT   and
          buhgalter-a   zamestitel-ju
          accountant-GEN   deputy-DAT

       c. poddakivanie   sekretar-em   director-u   i
          consenting   secretary-INSTR   director-DAT   and
          zamestitel-ju   buhgalter-om
          accountant-INSTR   deputy-DAT

       'consenting of the secretary to the director and of the accountant to the deputy'

The suggested analysis fits with the modelling of other argument supporting deverbal configurations, namely passive participles. In Gerasimova (to appear) the argument structure of passive participles was compared with that of nominalizations on the basis of three parameters: the position of arguments with regard to the head, the order of arguments ("Int" for internal; "Ext" for external), and the case assignment strategy.

	Case assignment strategy	Arguments always follow the head	The arguments' order
Nominalization type A	GEN	+	Ext Int
Nominalization type B	INSTR	+	Int Ext
Passive participles	INSTR	-	Ext Int

Table 1. Comparison of the two types of nominalizations with passive participles

This study has shown that the two types of nominalizations demonstrate similarities that differ from the structure of participles. As for both nominalizations with an external argument in GEN and passive participles speakers prefer the precedence of the external argument, whilst the external argument in INSTR follows the internal argument. On the contrary, the external argument with a passive participle is obligatorily assigned INSTR, as in the case of vP nominalization. This serves as evidence for the presence of vP in passive participles. However, the different word order of arguments in the case of nominalizations with INSTR probably shows that there may be verbal projections other than vP, where the internal argument moves. This speculation may be legitimate because, as was mentioned in Section 2, AspP is the maximum possible projection in Russian process nominalization. However, further work needs to be done to model the differences in argument order.

6. Conclusion

The evidence from this study suggests that differential case marking results not from differences in case assignment mechanisms but from two structural options that are available to some speakers. Therefore, in Russian the parameter of intralingual variation is not the case marking strategy but the amount of structure that is nominalized.

The findings of the study leave open the question as to why INSTR is significantly less acceptable with unergatives without a PP, although following the analysis, nominalizations from unergative stems can embed the *v*P. This question probably concerns a broader discussion on the fairness of comparison between production and acceptability experiments. Notably, it is interesting to find out to what range speakers are consistent in evaluating the constructions they produce as acceptable. Moreover, it needs to be determined what level of unacceptability verifies that the structure is ungrammatical (see (Knyazev 2017) about an acceptability decrease due to grammatical effects and its relation to absolute ungrammaticality). A comparison of the two methods and analysis of the differences in the results they provide is a vital issue for future research.

To sum up, in this paper I have addressed the issue of differential case marking in Russian process nominalizations. I conducted two linguistic experiments focusing on what cases speakers choose in the process of speech production and how they estimate the acceptability of cases when reading sentences. The results of the experiments proved that there are two case assignment strategies and that they coexist within the case grammar of one speaker. Finally, I suggested an analysis stating that the choice of case correlates with the structural shape of nominalization.

Anastasia Gerasimova
Lomonosov Moscow State University (MSU)
Moscow State University of Education (MSPU)
anastasiagerasimova432@gmail.com

References

Abney S. (1987). *The English noun phrase in its sentential aspect*. Ph. D. thesis, MIT, Cambridge (Mass.).

Alexiadou. A. (2001). *Functional Structure in Nominals: Nominalization and Ergativity.* J. Benjamins Publishing, Amsterdam.

Anderwald L. (2001). Was/were-variation in non-standard British English today. *English World-Wide* 22:1, pp. 1-21.

Baker M. C. (1985). *Incorporation: a theory of grammatical function changing*. Ph.D. thesis, MIT, Cambridge (Mass.).

Baker, M. C. & N. Vinokurova. (2010). Two modalities of case assignment: Case in Sakha. *Natural Language & Linguistic Theory* 28:3, pp. 593-642.

Chomsky, N. & H. Lasnik. (1993). The theory of principles and parameters. Jacobs, J., A. Stechow, W. Sternefeld & T. Vennemann (eds.), *Syntax: An international handbook of contemporary research,* Walter de Gruyter, Berlin, pp. 506-569.

Gerasimova A. (to appear). Differential case marking and positional distribution: Evidence from Russian event nominalizations. *Trudy Instituta lingvisticheskikh issledovanii* [*Works of the Institute for Linguistic Studies RAS*], Acta Linguistica Petropolitana, Saint Petersburg.

Grimshaw J. (1990). *Argument Structure*. The MIT Press, Cambridge.

Knyazev M. (2017). The effect of non-agentivity and constraints on čto-clauses in Russian: experimental study.

Paper presented at *The 5th International Conference «Russian speech culture»*, Moscow, February.

Koptjevskaja-Tamm M. (1993). *Nominalizations*. Routledge, London.

Koptjevskaja-Tamm M. (2013). Action Nominal Constructions. Dryer, M.S. & M. Haspelmath (eds.), *The World Atlas of Language Structures Online*, Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/62

Kratzer A. (1996). Severing the external argument from its verb. Rooryck, J. & L. Zaring (eds.), *Phrase structure and the lexicon*, Springer, Amsterdam, pp. 109-137.

Levin, T. & O.Preminger. (2015). Case in Sakha: are two modalities really necessary? *Natural Language & Linguistic Theory* 33:1, pp. 231-250.

Lyutikova E. (2014). Russian genitive possessor and formal models of the noun phrase. Lyutikova, E. A., A. V. Zimmerling & M. B. Konoshenko (eds.), *Tipologija Morfosintaksicheskih Parametrov. Materialy mezhdunarodnoj konferencii "TMP-2014"* [*Typology of Morphosyntactic Parameters 2014. Proceedings of The International Conference "TMP-2014"*], Vol. 1, MSPU, Moscow, pp. 120-145.

Marantz A. (1991). Case and licensing. G. Westphal, B. Ao & Chae H (eds.), *Proceedings of the 8th Eastern States Conference on Linguistics (ESCOL 8), Ithaca, NY*. Westphal, G., B. Ao & H. Chae (eds.), CLS Publications, pp. 234–253.

Woolford E. (2006). Lexical case, inherent case, and argument structure. *Linguistic inquiry* 37:1, pp. 111-130.

Pazel'skaya, A. & S. Tatevosov. (2008). Deverbal nouns and the structure of Russian verb. Plungyan, V.A. & S.G. Tatevosov (eds.), *Issledovaniya po glagol'noi derivatsii [Studies on verbal derivation],* Yazyki slavyanskikh kul'tur, Moscow, pp. 348-80.

Pereltsvaig, A., E. Lyutikova & A. Gerasimova. (2016). Case Marking in Russian Eventive Nominalizations: Inherent vs. Dependent Case Theory. Paper presented at *Formal Approaches to Slavic Linguistics 25*, Ithaca, NY, May.

Pereltsvaig, A., E. Lyutikova & A. Gerasimova. (forthcoming). Case Marking in Russian Eventive Nominalizations: Inherent vs. Dependent Case Theory.

Pesetsky D. (2003). *Principles & Parameters Theory*. Oxford International Encyclopedia of Linguistics.

Roeper T. (1987). Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18, pp. 267–310.

Shvedova N. (1980). *Russkaja Grammatika* [*Russian Grammar*]. AN SSSR Publ., Moscow.

Tatevosov S. (2008). Nominalization and the problem of indirect access. Bondarko, A.V., G.I. Kustova & R.I. Rozina (eds.), *Dinamicheskiie modeli. Slovo. Predlozhenie. Tekst. Sbornik statei v chest' E.V Paduchevoi.* [*Dynamic models. Word. Sentence. Text. Collected works in honor of E.V. Paducheva*], Yazyki slavyanskikh kul'tur, Moscow, pp. 750-773.

Hill Mari verbal constructions with *ǝl'ǝ / ǝlǝn*: discontinuous past and beyond[1]

Daria Mordashova

This paper deals with Hill Mari verbal constructions with *ǝl'ǝ/ǝlǝn*, verbal clitics, which are etymologically 'frozen' past tense forms of the verb 'to be'. The clitics are considered to be discontinuous past markers. Applied to full-fledged verbs, these markers follow polysemy patterns typical of the discontinuous past in other languages. Nevertheless, they also function within the scope of mirativity, and such polyfunctionality has not yet been mentioned or accounted for in typology. The possible reasons for such a semantic development of discontinuous past markers are discussed in the paper.

1. Introduction

This research deals with analytical verbal constructions with *ǝl'ǝ/ǝlǝn* (verbal clitics, which are etymologically 'frozen' AOR.3SG / PRF.3SG forms of the verb 'to be') in Hill Mari (Mari < Finno-Permic < Finno-Ugric < Uralic). This language is mainly spread across the Gornomari district of Mari El in Russia.

The data used in this paper were collected during a field trip to Kuznetsovo village and the nearby villages of Tyumanovo, Malyj and Bolshoj Apshak-Pelyak, Paulkino and Kukshilidy in 2016. The main source of information was elicitation (such examples will be provided in this paper without any special notes). Furthermore, corpus data were used, and examples from this source have a note 'corpus' added in brackets.[2]

The aim of this paper is to investigate the structure and semantics of constructions with *ǝl'ǝ/ǝlǝn* in Hill Mari. Firstly, in Section 2 the tense and aspect system of Hill Mari is described, concentrating on the analytical forms. Section 3 presents both previous study of the problem and the new Hill Mari data regarding what meanings can be expressed by the constructions in question. In Section 4 the analysis of the verbal constructions with clitics is proposed, in terms of both semantics and distribution of clitics. Conclusions are drawn in Section 5.

2. Tense and aspect system of Hill Mari

The Hill Mari tense system is characterized by a binomial opposition of non-past versus past verbal forms. It includes no formal distinction between present and future (1), and the non-past verbal form can be used in progressive, habitual and generic contexts (following the TMA questionnaire proposed in Dahl 1985). The inventory of forms with reference to the past is much broader: it contains two synthetic past tenses (aorist and perfect[3], see (2)-(3); these labels are explained below) and several analytical forms with clitics, which will be analyzed in this paper.

(1) *tǝdǝ sirmäš-ǝm **sir-ä**.*
    he   letter-ACC   write-NPST.3SG
     1. 'He <u>is writing</u> a letter.'
     2. 'He <u>will write</u> a letter.'

(2) *tǝdǝ sirmäš-ǝm **sir-ǝš**.*
    he   letter-ACC   write-AOR.3SG
    'He <u>wrote</u> a letter.'

(3) *tǝdǝ sirmäš-ǝm **sir-en**.*
    he   letter-ACC   write-PRF.3SG
    'He wrote a letter.'

Mari grammars (see Pengitov 1961; Savatkova 2002) have traditionally claimed that the distinction between the aorist and the perfect concerns evidentiality. The aorist was claimed to mark firsthand evidence, while the perfect to express non-firsthand evidence, including all the possible indirect sources of information (inferential, reportative etc., see Aikhenvald 2004). As it is described in Alhoniemi (1993:115), the perfect in Mari is the tense of indirect description and non-firsthand observation.

   However, the topic is revisited in Golosov & Kozlov (2017), where alternative factors are taken into consideration. It is argued that the choice between the aorist and the perfect in Hill Mari is determined by the discourse structure and the notion of *perspective* (see Kamp & Reyle 1993). Kamp & Reyle 1993 propose the term *perspective time* in their analysis of 'extended flashbacks' like (4).

(4) *Fred arrived at 10. He <u>had got up</u> at 5; he <u>had taken</u> a long shower and <u>had left</u> the house at 6:30.*

The perspective time in (4) is set by Fred's arrival, and the past perfect encodes the precedence relation between the events and perspective time. In the case of the past simple, perspective time coincides with the utterance time. Thus, the past simple tense simply marks the events preceding the utterance time.

   In Paducheva (2010 [1996]) this theory is developed with regard to the phenomenon of narrative, as the hypotheses concerning simple past in Kamp & Reyle 1993 refer only to dialogical contexts. In particular, it is claimed that perspective time (or *position of observer* – a notion that is more common in the Russian academic tradition) coincides with the time of events in the narrative (the so-called 'story-now' of the narrative). That is the reason why verbal forms may have different interpretations depending on the type of discourse (i.e.

---

[3] It should be noted that these tense labels are rather conventional and may not correspond to what is referred to as the aorist and perfect in other languages.

dialogue vs. narrative). For instance, the Russian sentence in (5) can have two different interpretations (see Paducheva 2010:12).

(5) *На   стен-е   справа   висе-л-а   картин-а.*
   on   wall-LOC   on.the.right hang-PST-F.SG   painting-NOM
   1) 'A painting <u>hung</u> on the wall on the right (<u>some time ago</u>, but it is no longer on the wall).'
   2) 'A painting <u>was hanging</u> on the wall on the right (<u>at that very moment</u>; nothing is said about the current state of affairs).'

The first interpretation appears in the context of a dialogue, while the second is possible only within a narrative.

   Golosov & Kozlov (2017) apply the concept of perspective to the semantics of Hill Mari past tenses. The aorist in Hill Mari is specific in that it is a tense demanding coincidence of event time and perspective time. It means that the aorist is possible in a dialogical context if the event time precedes the utterance time (which coincides with the perspective time). In a classical narrative, the aorist needs the event time to coincide with the 'story-now' of the narrative (i.e. the perspective time). The perfect, according to Golosov & Kozlov (2017), on the contrary, tends to be insensitive to perspective and is simply an unmarked past tense. That is why the perfect is used as a default past tense in Hill Mari.

   By way of proof, a narrative from Golosov & Kozlov (2017) is cited here in (6). The perfect is exploited to open the narrative and set the perspective time in the past, and after that the aorist and perfect can equally be used to mark a chain of events.

(6)   a. *ikänä barin   **keles-en**   što   t'ot'a-m   saltak-ôš*
       once landowner say-PRF.3SG that grandfather-POSS.1SG soldier-ILL
      ke-šäšlôk.
      go-PTCP.DEB
  b. *môn'-ôn kogo papa   **tol-ôn**   / **tol'-ô**   tôdä-n   dokô*
      I-GEN   big grandmother go-PRF.3SG go-AOR.3SG that-GEN   to
      *dä **keles-en**   / **keles-ôš**   što   värô-šô-žô   môn' ške=ok*
      and say-PRF.3SG   say-AOR.3SG that place-ILL-POSS.1SG I   self=EMPH
      saltak-ôš   ke-em.
      soldier-ILL go-NPST.1SG
  c. *barin   kogon   **ör-ôn**   / **ör'-ô**.*
      landowner very   be.surprised-PRF.3SG be.surprised-AOR.3SG
   '{My grandfather was lame in the right leg.} Once the landowner said_{prf/*aor} that my grandfather had to join the army. My great-grandmother came_{aor/prf} to the landowner and said_{aor/prf} that she would join the army instead of him. The landowner got_{aor/prf} very much surprised.'   (Golosov & Kozlov 2017:17-18)

I will return to the distinction between aorist and perfect in Section 4, but now let us switch topic and return to the analytical constructions with *ôl'ô/ôlôn*, which are the main topic of interest in this paper. The marker *ôl'ô/ôlôn* (etymologically 'frozen' AOR.3SG/PRF.3SG forms of the verb 'to be') combines with verbs in both non-past and past tenses (although only in the perfect), following the rules of agreement in person and number with the subject. As can be seen in (7)-(8), the clitics themselves do not agree with the subject.

(7) *mǝn' šǝšer-ǝm* **jü-äm**   *ǝl'-ǝ*   / *ǝl-ǝn*   / **ǝl'-ǝm*
     I   milk-ACC   drink-NPST.1SG be-AOR.3SG   / be-PRF.3SG   / be-AOR.1SG
     / **ǝl-ǝn-am.*
     / be-PRF-1SG
     'I <u>drank</u> milk.'

(8) *tǝn' möskä-m* **pušt-ǝn-at**   *ǝl'-ǝ*   / *ǝl-ǝn*   / **ǝl'-ǝc*
     you bear-ACC   kill-PRF.2SG   be-AOR.3SG   / be-PRF.3SG   / be-AOR.2SG
     / **ǝl-ǝn-at.*
     / be-PRF-2SG
     'You <u>killed</u> a bear.'

It should be noted that Kangasmaa-Minn (1960) points out on the agreement of clitics with a plural subject in Meadow Mari, although Hill Mari does not follow such a pattern, as we can observe in (9).

(9) *mä   ängǝr-ǝš kašt-aš* **jarat-en-nä** *ǝl'-ǝ*   / *ǝl-na
     we   river-ILL go-INF   love-PRF.1PL be-AOR.3SG   / be-AOR.1PL
     / *ǝl-ǝn*   / **ǝl-ǝn-na.*
     / be-PRF.3SG   / be-PRF-1PL
      'We <u>loved</u> going to the river.'

3.   *Semantics of the constructions on ǝl'ǝ / ǝlǝn*

Previous works containing Hill Mari data regarded constructions on *ǝl'ǝ/ǝlǝn* as separate verbal tenses. For instance, the grammar of Mari (see Alhoniemi 1993) considers these forms as a periphrastic imperfect (for non-past tense + clitics) and a periphrastic perfect (for perfect + clitics), but does not pay much attention to their grammatical semantics. A similar view is found in Pengitov (1961), Serebrennikov (1960) and Savatkova (2002), where labels such as 'prolonged past' and 'pluperfect' are attached to the constructions in question.

   In this paper I argue that it is unnecessary to maintain such an extended system of tenses in Hill Mari. The clitics *ǝl'ǝ/ǝlǝn* can be considered as independent markers, which function within the realm of the *discontinuous past*. According to Plungian and van der Auwera (2006:317), the discontinuous past is a type of past time reference 'with no present relevance'. Markers of this category apply not to verbal stems (as classical tense markers do), but rather to full-fledged verbal forms. If typical markers of tense set the temporal reference of the verb on the time axis, discontinuous past markers only *modify* the default temporal interpretation of the verbal form and 'move' the situation backwards on the time axis (see (10)-(11)). As it is stated in Plungian and van der Auwera (2006:344), 'this change amounts in introducing a temporal (or notional) break between the point of reference and the situation: metaphorically, one can speak about a kind of "detachment" or "shift"'. That is why in the Russian academic tradition (cf. Plungian 2001, Sitchinava 2013) they acquire the definition of 'retrospective shift' markers (abbreviations RETR1 and RETR2 will appear under *ǝl'ǝ* and *ǝlǝn* further on).

(10) *mǝn' škol-ǝštǝ*   **päšäl-em**.
     I   school-IN   work-NPST.1SG
     'I <u>work</u> at school.'

(11) *mə̂n' škol-ə̂štə̂*   ***päšäl-em***   ***ə̂l'-ə̂ / ə̂l-ə̂n***.
      I   school-IN   work-NPST.1SG   RETR1 / RETR2
      'I <u>worked</u> at school.'

In combination with full-fledged verbal forms Hill Mari clitics *ə̂l'ə̂/ə̂lə̂n* have developed a range of meanings, which is set forth below.

3.1. Discontinuous past
3.1.1. Framepast

Following Plungian and van der Auwera (2006), the label 'discontinuous past' is used as a cover term for imperfective- and perfective-based values, such as habitual, progressive or pluperfect: "with imperfective verbs, the markers of discontinuous past encode situations of limited duration, which are claimed not to extend up to the moment of speech" (Plungian & van der Auwera 2006:323). The meaning of the past habitual can be expressed by the combination of the clitics *ə̂l'ə̂/ə̂lə̂n* with finite verbs in both the non-past (12)-(13) and the perfect (14)-(15).

(12) *pet'a izi-žə̂*   *god-ə̂m každə̂j kečə̂-n lem-ə̈m*   ***kačk-eš***
      P.   small-POSS.3SG while   every   day-GEN soup-ACC   eat-NPST.3SG
      ə̂l'ə̂ / ə̂lə̂n.
      RETR1 / RETR2
      'Pet'a <u>ate</u> soup every day in his childhood.'

(13) *ärvätə̂-län-žə̂*   ***čikt-ät***   *oš*   *tə̂gə̂r-ə̂m, šarpan-ə̂m*
      bride-DAT-POSS.3SG dress-NPST.3PL   white   shirt-ACC   sharpan-ACC
      pə̂nz-en   ***šə̂nd-ät***   ***ə̂lə̂n***.
      dress.up-CVB   put-NPST.3PL   RETR2
      '(The family) <u>dressed</u> the bride into a white shirt and <u>put</u> a sharpan <u>on</u> her.' (corpus; the text describes wedding traditions, and sharpan is a typical Mari decoration for the bride)

(14) *mä*   ***šə̂nz-en-nä***   ***ə̂l'ə̂ / ə̂lə̂n***   *okn'a* *anzə̂l-nə̂*   *dä*
      we   sit-PRF-1PL   RETR1 / RETR2   window before-IN2   and
      anž-en-nä   *lə̂m-ə̂m*.
      look-PRF-1PL   snow-ACC
      'It often happened that we <u>sat</u> in front of the window and <u>looked</u> at the snow.'

(15) *mä*   *kečə̂väl šo-mə̂-kə̂*   *kačk-aš* ***näng-en-nä***   ***ə̂lə̂n***.
      we   noon   reach-PTCP.PASS/NMLZ-ILL2 eat-INF bring-PRF-1PL RETR2
      'We <u>brought</u> the food at noon.' (corpus; the text describes a typical working day in a village)

The examples above report about repeated situations in the past, and there is no information concerning the current relevance of the situations. It could be assumed that the event expressed by the past habitual no longer holds at present, although this semantic component is not compulsory, as we can see in (16).

(16) *molnam mǟn'ǝ šǟre-n=ok   lâdâš-vlä-m **sir-en-äm***
     before I   frequent-GEN=EMPH poem-PL-ACC write-PRF-1SG
     ***âl'ǝ / âlân**, dä   kǝzǝt=ät   sir-em.*
     RETR1 / RETR2 and   now=ADD   write-NPST.1SG
     'I often <u>wrote</u> poems before, and I still do that.'

As for the choice of finite verbal forms with clitics, the non-past and perfect tenses are usually interchangeable in past habitual contexts. Used independently (without clitics), both these tenses can be interpreted habitually (Shcherbinin 2016, see (17)-(18)). Clitics in such contexts do not change the already existing temporal and aspectual features of the situation; they only 'move' it backwards on the time axis.

(17) *vsegda **kǝn'ǝl-äm**   irok   kud cäš-ǝn.*
     always get.up-NPST.1SG morning six   hour-GEN
     'I always <u>get up</u> at 6 o'clock in the morning.'

(18) *tǝdǝ sirmäš-vlä-m **sir-en**.*
     he   letter-PL-ACC write-PRF.3SG
     '(What did your brother usually do after breakfast?) He <u>wrote</u> letters.'

The combination of clitics with verbs in the non-past may also express the meaning of past progressive (19)-(20) to mark an 'on-going activity' or a background action in case of two simultaneous actions. In progressive contexts the clitic *âl'ǝ* tends to be far more preferable than *âlân*.

(19) *kânam mǟn' kǝdež-ǝš   pâr-ǝš-ǝm   ävä-m   telefon dono*
     when   I   room-ILL   enter-AOR-1SG mother-POSS.1SG phone   with
     ***pop-a   âl'ǝ / ²²âlân**.*
     talk-NPST.3SG RETR1 / RETR2
     'When I entered the room, my mother <u>was talking</u> on the phone.'

(20) *ävä-štǝ   tol-mâkǝ, t'et'ä-vlä **amal-at**   âl'ǝ / ²²âlân.*
     mother-POSS.3PL come-CVB child-PL   sleep-NPST.3PL RETR1 / RETR2
     'When mother came, children <u>were</u> already <u>sleeping</u>.'

The replacement of non-past finite forms in (19)-(20) by perfect results in an automatic pluperfect interpretation, with a standard meaning of precedence in the past (cf. (21)-(22), which differ only in the choice of the tense of the verb *go down* and, thus, also the interpretation of the main clause as durative or completive).

(21) *kânam mä tol-ân-na,   kečǝ **šǝnz-eš**   âlân / âl'ǝ.*
     when   we come-PRF-1PL sun go.down-NPST.3SG   RETR2 / RETR1
     'When we returned, the sun <u>was going down</u>.'

(22) *kânam mä tol-ân-na,   kečǝ **šǝnz-ǝn**   âlân / ²²âl'ǝ.*
     when   we come-PRF-1PL sun go.down-PRF.3SG   RETR2 / RETR1
     'When we returned, the sun <u>had</u> already <u>gone down</u>.'

In sentences such as (22) the use of the clitic *∂l'∂* is considered by Hill Mari consultants to be less felicitous. The distribution of clitics may be determined by the functional difference between the aorist (*∂l'∂*) and the perfect (*âlân*). The problem of how the two clitics are distributed is discussed in Section 4.

3.1.2. 'Cancelled' result

As discontinuous past markers, *∂l'∂/âlân* also denote that the consequent state or the result of an action do not exist at the moment of speech (see also the notions of 'cancelled'/'reversed' result in Squartini 1999 or 'anti-resultative' in Plungian 2001). This interpretation occurs when the clitics are combined with verbs in the perfect (23)-(24), especially with telic verbs having a clear resultative component.

(23) *m∂n' bilet-∂m* **näl-en-äm** **∂l'∂ / âlân,** *no vara t∂d∂-m*
I　　ticket-ACC take-PRF-1SG　　RETR1 / RETR2 but　then　it-ACC
sdaj-∂š-em.
return-AOR-1SG
'I <u>had bought</u> a train ticket, but then returned it.'

(24) *van'a　šärg∂-št∂* **jam-ân** **∂l'∂ / âlân** *vara t∂d∂-m*
V.　　　forest-IN　　get.lost-PRF.3SG　RETR1 / RETR2 then　he-ACC
mo-n-∂t.
find-PRF-3PL
'Vanya <u>got lost</u> in the forest, but was found then.'

It should be noted that the scope of discontinuous past markers includes not a situation itself, but its consequent phase, namely "the only available temporal fragment which can have some duration and be extendable to the present" (Plungian & van der Auwera 2006:324).

Nevertheless, the clitics can also mark a situation in which a resulting state has not been attained, as the action has been interrupted in the medial phase (25). Arkadiev (2014) describes such a situation in terms of an 'unsuccessful attempt'.

(25) *xala-šk∂　ke-äš* **tärvän-en-nä** **∂l'∂,** *no　pel　korn∂-št∂　m∂ngeš*
town-ILL　go-INF　set.off-PRF-1PL　RETR1　but　half　road-IN　　back
särnäl-nä.
return-AOR.1PL
'We <u>set off</u> for the town, but turned back halfway.'

However, the interpretation of 'unattained result' does not appear in contexts, when the resulting state of the verb seems to be difficult to 'cancel', as in (26).

(26) *Kot'i* **kol-en** **∂l'∂,** *no　veterinar　veremä-št∂ ukol-âm*
cat　die-PRF.3SG　RETR1　but　vet　　　　time-IN　　injection-ACC
∂št-en.
do-PRF.3SG
*'The cat <u>nearly died</u>, but the vet gave it an injection right on time.'
OK'The cat <u>died</u>, but the vet gave it an injection (and the cat came back to life).'

3.2. Mirativity

Probably the most typologically challenging data are those on the mirative semantics of the constructions with *âl'â/âlân*. The notion of 'mirativity' was introduced by Scott DeLancey, and it stands for 'the linguistic marking of an utterance as conveying information which is new or unexpected to the speaker' (see DeLancey 1997, 2001:369-370). The mirative extension of discontinuous past markers has not yet received a great deal of attention in typology. However, the use of the pluperfect encoding evidentiality and mirativity (which is closely related to evidentiality) is typologically attested: for instance, it takes place in colloquial German, as is argued in Hennig (2000) and Ammann (2005), as well as in the Bezhta language (see Testelets & Halilov 2002).

The existing descriptions of languages closely related to Hill Mari do not mention any mirative uses of discontinuous past markers (cf. Kolyadenkov 1962 on Moksha and Erzya, Tsypanov 2005 on Komi-Zyrian, Sitchinava 2013 on Udmurt and Meadow Mari, Kiss 2017 on Old Hungarian). For instance, the Udmurt language with its particle *val* is claimed to 'represent a system with a maximally "broad" discontinuous past marker' (Plungian & van der Auwera 2006:342). It develops uses within the domains of framepast and 'cancelled' result (see Sitchinava 2013: 200-201) and displays a kind of an attenuative meaning when applied to imperative. Nevertheless, it is not stated whether the mirative reading is available for this particle.

In Hill Mari the mirative meaning appears when discontinuous past markers apply to non-past verbal forms which have a progressive or habitual reading. It is especially productive in dialogical contexts with reference to the present such as in (27)-(28), but it is also found in published folklore texts (29).

(27) a. *van'a,   tän' ma   piš   sâlâk   âl-at?*
          V.   you   what very sad   be-NPST.2SG
      b. *dä   anž-et   gân', pet'a zabor-âm   **čiält-ä**   **âlân**/ ⁇*âl'â!*
          well look-NPST.2SG   PTCL P.   fence-ACC paint-NPST.3SG   RETR2/RETR1
      c. *a   vet   mäen' äške čiält-än-em   âl'â.*
          but PTCL I   self paint-OPT-1SG   RETR1
      'Vanya, why are you so sad?
       Well, look, Petya <u>is painting</u> the fence! But I wanted to do it myself.'

(28) a. *mäen' tengečâ   orkestr-âštâ   vâstupaj-en-äm.*
          I   yesterday orchestra-IN   perform-PRF-1SG
      b. *cüdä   vele, tän' eče orkestr-âšt-at   **šakt-et**   **âlân**/⁇*âl'â!*
          miracle only you even orchestra-IN-ADD play-NPST.2SG   RETR2/RETR1
      'I had a performance with an orchestra yesterday.
       What a miracle, you even <u>play</u> in an orchestra!'

(29) *pišok patâr   **âl-at**   **âlân-âš.***
       very strong   be-NPST.2SG   RETR2-EMPH
       'It appears that you <u>are</u> really strong.'   (Aktsorin 1995:22)

Mirative semantics can also occur in contexts referring to the past. In this case discontinuous past markers combine with verbs in the perfect, frequently with a first person subject and a lack of agentive control (compare (30) and (31)).

(30) *mə̂n' tengečə̂   omə̂n-ə̂stə̂   **kašt-ə̂n-am**   (ə̂lə̂n/ʔʔə̂l'ə̂).*
     I   yesterday   sleep-IN   walk-PRF-1SG RETR2/RETR1
     'It turned out that I <u>walked</u> while sleeping yesterday.'

(31) *mə̂n' tengečə̂   lapka-š   **kašt-ə̂n-am**   (ʔə̂l'ə̂/*ə̂lə̂n).*
     I   yesterday   shop-ILL   walk-PRF-1SG RETR1/RETR2
     'I <u>went</u> to the shop yesterday.'

Strictly speaking, it is possible to omit *ə̂lə̂n* in contexts such as (30), but it is impossible to use this marker in (31), as it would indicate, according to my consultants, that the speaker *does not remember* how the action took place (such implications on the lack of control are typical of evidential markers, see Aikhenvald 2004).

   This fact may lead to a hypothesis concerning the development of mirative semantics of *ə̂l'ə̂/ə̂lə̂n* in Hill Mari as an extension of evidentiality. It was mentioned in Section 2 that traditional Mari grammars explain the difference between the aorist and perfect tenses in terms of evidentiality: the perfect marks non-firsthand evidence, while the aorist marks firsthand evidence (see Savatkova 2002:199-203). It should be borne in mind that *ə̂lə̂n* is a 'frozen' perfect form, which is more preferable in mirative contexts. Thus, it seems rather logical to assume that the discontinuous past marker in the form of the perfect tense is chosen to express evidentiality, which develops into mirativity in a certain range of contexts (such semantic development is widely attested typologically, see Aikhenvald 2004, 2012).

   However, the marker *ə̂l'ə̂* (in the form of the aorist, which is claimed in grammars to encode firsthand evidence) is also compatible with the mirative meaning. Although it is less preferable for my consultants, no strict constraints on the use of *ə̂l'ə̂* have been found so far. Moreover, 'evidential' differences within the type of evidence (firsthand vs. non-firsthand) are no longer maintained in Hill Mari: see (32)-(34), where the aorist and the perfect are interchangeable in contexts with direct as opposed to indirect sources of information.

(32) *lapka-š sə̂kə̂r-ə̂m **kand-en-ə̂t** / ᴼᴷ**kand-evə̂,**   mə̂n' ške=ok*
     shop-ILL bread-ACC bring-PRF-3PL bring-AOR.3PL   I   self=EMPH
     tidə̂-m   už-ə̂n-am.
     that-ACC   see-PRF-1SG
     'They <u>brought</u> bread to the shop, I saw it myself.' (direct visual evidence)

(33) *kə̂dal   ert-ə̂šə̂   mašinä juk-ə̂m   kol'-ə̂m,*
     drive.CVB   go-PTCP.ACT   car   voice-ACC   hear-AOR.1SG
     *lapka-š sə̂kə̂r-ə̂m **kand-evə̂**   /ᴼᴷ**kand-en-ə̂t** vekät.*
     shop-ILL bread-ACC bring-AOR.3PL   bring-PRF-3PL   probably
     'I heard a car passing by, it seems like they <u>brought</u> bread to the shop.' (non-visual evidence)

(34) *lapka-š sə̂kə̂r-ə̂m **kand-en-ə̂t** /ᴼᴷ**kand-evə̂**   man-ə̂t.*
     shop-ILL bread-ACC bring-PRF-3PL bring-AOR.3PL   say-NPST.3PL
     tidə̂ gišän   paškud-em   keles-ə̂š.
     that about   neighbour-POSS.1SG tell-AOR.3SG
     'It is said that they <u>brought</u> bread to the shop. My neighbour told me that.' (reportative)

All of these factors weaken the hypothesis that mirativity developed from evidentiality in Hill Mari. In response to this problem, I propose another scenario, which is elaborated in the following section.

4. *Discussion*

Section 3 shed the light on the semantics of Hill Mari constructions with *əl'ə̂/ələ̂n*, but the question as to where such semantics actually comes from still remains unanswered. In previous studies discontinuous past markers were analyzed as polysemantic and polyfunctional units. Thus, Sitchinava (2009, 2013:232) does not aim to propose any semantic invariant for the Russian particle *bylo*, only providing a broad list of its meanings, such as framepast and 'cancelled' result. The idea of a possible invariant was hinted in Plungian and van der Auwera (2006:346), but was not clearly elaborated. To the contrary, I argue that all the examples described above do not form a bundle of separate meanings specific for each combination of a full-fledged verb and a clitic itself. Different meanings appear only in context, when a particular interpretation of the verbal form interacts with the semantics of the discontinuous past markers.

For instance, non-past verbal forms can have progressive or habitual readings. Consequently, when such forms combine with discontinuous past markers, they undergo a 'transformation' into the past progressive or past habitual. The perfect in Hill Mari, depending on the context, can encode a habitual situation or express a completed situation and its resulting state. The attachment of discontinuous past markers results in the meanings of past habitual, anteriority or 'cancelled' result.

I surmise that all these semantic 'transformations' are dominated by a common semantic invariant of discontinuous past markers, which is closely related to the notion of *distance*. Recall that the clitics *əl'ə̂/ələ̂n* do not modify the existing interpretation of the verb in context (its temporal and aspectual characteristics), but introduce a temporal break (distance) between the point of reference and the situation, carrying out a kind of 'shift' backwards on the time axis. The semantics of 'cancelled' result appears when the context implies that the result of an action is currently irrelevant. This can also be viewed in terms of distance or break between two adjacent states of a situation (the result state and the new state when the result is no longer maintained). In this paper I show that the proposed invariant (the notion of distance) is also applicable to mirative contexts. Thus, Hill Mari data provide a new extension for the semantics and typology of the discontinuous past.

The role of distance is emphasized as being crucially important when explaining the nature of such categories as evidentiality and mirativity. Evidential markers usually provide "self-distancing of the speaker from the situation, about which surprise is being expressed" (Chirikba 2003:264). I would assume that the same effect takes place in Hill Mari mirative contexts. In this case, one is dealing with the semantic development of discontinuous past markers, when the temporal distance is reinterpreted metaphorically. Thus, what is introduced here is a notional break between the speaker's expectations and her realization of an actual situation. This phenomenon has not yet been investigated from a typological perspective, so the conclusions drawn in this paper could make a contribution to the typology of both discontinuous past and mirative markers.

As for the latter, mirative meanings are well-known to develop as extensions of evidentials (cf. evidentials as mirativity strategies in Aikhenvald 2004:207-209; earlier highlighted by DeLancey 1997, 2001) or of other categories, such as perfect, irrealis or hypothetical mood.

Formally, the range of mirative meanings can be expressed through dedicated complex verbal constructions (involving the verbs *be*, *become*, *discover* or a grammaticalized copula), verbal affixes, particles, or through a special series of pronouns (see Aikhenvald 2012:438-458). However, the verb *be* in the constructions from Aikhenvald's inventory has not been claimed to function beyond the scope of mirativity and evidentiality, whereas the Hill Mari markers analyzed in this paper are highly polyfunctional. In this regard, Hill Mari provides new data on the typology of mirative markers derived from *be*-predicates.

Another question that needs addressing concerns the distribution of clitics. As we saw in Section 3, in some contexts, such as the past habitual and 'cancelled' result, clitics are freely interchangeable, while in other contexts (e.g. past progressive, pluperfect and mirative) one of the clitics tends to be far more preferable. In Voronov & Mordashova (2016) the first data on their distribution are provided (see Table 1).

Semantics	*âl'ə̂*	*âlə̂n*	Example
NPST + clitics			
Past progressive	OK	??	(19)-(20)
Past habitual	OK	OK	(12)-(13)
Mirativity	??	OK	(27)-(29)
PRF + clitics			
Pluperfect	??	OK	(22)
Past habitual	OK	OK	(14)-(15)
'Cancelled' result	OK	OK	(23)-(24)
Mirativity	??	OK (provided lack of agentive control)	(30)

Table 1: The distribution of clitics in different contexts

In Section 3.1 I proposed that the distribution of *âl'ə̂/âlə̂n* may be determined by the functional difference between the aorist (*âl'ə̂*) and the perfect (*âlə̂n*). As is argued in Golosov & Kozlov (2017), the aorist demands a coincidence of the event time with the perspective time. This fact can explain why the marker *âl'ə̂* is preferred in progressive contexts rather than *âlə̂n*: perspective time in contexts such as (19)-(20), which is set by the subordinate clause (e. g. *when mother came*), coincides with the event time (e. g. *children were sleeping*). In pluperfect contexts, on the contrary, such a coincidence is hardly possible, as the precedence of the event time is implied by the definition of the pluperfect itself. This leads to a more frequent choice of *âlə̂n* (the perfect form).

The explanation of a higher preference for *âlə̂n* in mirative contexts tends to be the most challenging. The notion of distance may again be helpful, as mirative markers let the speaker 'distance' herself from an unexpected situation. The form of the aorist with its 'short distance' effect, which occurs due to the coincidence of the event time and the perspective time, may be less suitable in mirative contexts.

5. *Conclusion*

This research provides new typological data on the functions of discontinuous past markers. The main topic of interest here are verbal constructions with *âl'ə̂/âlə̂n* in Hill Mari, which have previously been described as independent tense forms. In my paper it is shown that the clitics *âl'ə̂/âlə̂n* should rather be viewed as separate markers within the framework of the

discontinuous past. The attachment of these markers to full-fledged verbs develops both typologically expected meanings from framepast and 'anti-resultative' clusters and rather challenging mirative semantics. As far as the distribution of clitics is concerned, it seemingly depends on the functional difference between the aorist and the perfect, which is related to the arrangement of the event time and the perspective time.

The meanings of the constructions in question follow a semantic invariant for discontinuous past markers, which has not been elaborated on in previous studies, except for some initial contributions in Plungian and van der Auwera (2006). Their research does not, however, involve mirativity. The essence of the proposed invariant lies in the notion of distance. The temporal distance is introduced by the clitics as discontinuous past markers by definition, and the metaphorical distance occurs in mirative contexts. Thus, the Hill Mari data contribute to further research on the sources of mirative markers (see Aikhenvald 2012, Mexas 2016).

Abbreviations

1, 2, 3 – 1, 2, 3 person
ACC – accusative
ADD – additive particle
AOR – aorist
CVB – converb
DAT – dative
EMPH – emphatic particle
F – feminine
GEN – genitive
ILL – illative
ILL2 – 2nd illative
IN – inessive
IN2 – 2nd inessive
INF – infinitive
LOC – locative
MULT – multiplicative

NEG – negation
NMLZ – nominalization
NOM – nominative
NPST – non-past tense
OPT – optative
PL – plural
POSS – possessive
PRF – perfect
PST – past tense
PTCL – particle
PTCP.ACT – participle (active voice)
PTCP.DEB – participle (debitive mood)
PTCP.PASS – participle (passive voice)
RETR1 / RETR2 – discontinuous past markers
SG – singular

Daria Mordashova
Lomonosov Moscow State University
mordashova.d@yandex.ru
https://istina.msu.ru/profile/mordashova.d/

References

Aikhenvald, A. (2004). *Evidentiality.* Oxford University Press, Oxford.
Aikhenvald, A. (2012). The essence of mirativity. *Linguistic Typology* 16, pp. 435-485.
Aktsorin, V. (1995). *Skazki gornykh mari: Svod marijskogo fol'klora.* [Hill Mari fairytales: Collection of Mari folklore.] Marijskoye knizhnoye izdatel'stvo, Yoshkar-Ola.
Alhoniemi, A. (1993). *Grammatik des Tscheremissischen (Mari): mit Texten und Glossar.* Buske, Hamburg.
Ammann, A. (2005). Abbau und Anschwemmung: Doppelte Perfektformen und Grammatikalisierung im deutschen Tempussystem. Mortelmans, T. & T. Leuschner (eds.), *Grammatikalisierung im Deutschen,* Mouton de Gruyter, Berlin-N. Y., pp. 251-276.

Arkadiev, P. (2014). Sistema form plyuskvamperfekta v shapsugskom dialekte adygejskogo yazyka. [The system of pluperfect forms in the Shapsug dialect of Adyghe.] *Voprosy yazykoznaniya* 4, pp. 46-65.

Chirikba, V. (2003). Evidential category and evidential strategy in Abkhaz. Aikhenvald, A. & R. M. W. Dixon (eds.), *Studies in evidentiality*, John Benjamins Publishing Company, Amsterdam / Philadelphia, pp. 243-272.

Dahl, Ö. (1985). *Tense and aspect systems.* Blackwell, Oxford.

DeLancey, S. (1997). Mirativity: The grammatical marking of unexpected information. *Linguistic Typology* 1:1, pp. 33-52.

DeLancey, S. (2001). The mirative and evidentiality. *Journal of pragmatics* 33:3, pp. 369-382.

Golosov, F. & A. Kozlov (2017). Nablyudatel' v sisteme proshedshikh vremyon gornomarijskogo yazyka. [The position of Observer in Hill Mari past tenses.] In press. *Acta Linguistica Petropolitana.*
https://www.academia.edu/31668516/Наблюдатель_в_системе_прошедших_времён_горномарийского_языка_1

Hennig, M. (2000). *Tempus und Temporalität in geschrieben und gesprochenen Texten.* Niemeyer, Tübingen.

Kamp, H. & U. Reyle (1993). *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory.* Kluwer Academic Publishers, Dordrecht.

Kangasmaa-Minn, E. (1960). The so-called past tenses in Cheremis. *American studies in Uralic linguistics* 1, pp. 93-120.

Kiss, K. (2017) The rise and fall of Hungarian complex tenses. Mathieu E. & R. Truswell (eds.), *Micro-change and macro-change in diachronic syntax.* Oxford University Press, pp. 61-74.

Kolyadenkov, M. (1962). *Grammatika mordovskikh (mokshanskogo i erzyanskogo) yazykov.* [The grammar of Mordvin (Moksha and Erzya) languages.] Mordovskoye knizhnoye izdatel'stvo, Saransk.

Mexas, H. (2016). Mirativity as realization marking: A cross-linguistic study. [MA thesis]. Universiteit Leiden.
https://www.academia.edu/30875827/Mirativity_as_realization_marking_A_cross-linguistic_study

Paducheva, E. (2010, 1st edition 1996). *Semanticheskiye issledovaniya: Semantika vremeni i vida v russkom yazyke. Semantika narrativa.* [Studies in semantics: Semantics of tense and aspect in Russian. Semantics of narrative.] Yazyki russkoj kul'tury, Moscow.

Pengitov, N. (1961). *Sovremennyj marijskij yazyk. Chast' 2: Morfologiya.* [Modern Mari. Part 2: Morphology.] Marijskoye knizhnoye izdatel'stvo, Yoshkar-Ola.

Plungian, V. (2001). Antirezul'tativ: do i posle rezul'tata. [Anti-resultative: before and after result]. Plungian, V. (ed.), *Issledovanija po teorii grammatiki. I: Glagol'nyje kategorii,* Russkie slovari, Moscow, pp. 50-88.

Plungian, V. & J. van der Auwera (2006). Towards a typology of discontinuous past marking. *Sprachtypologie und Universalienforschung – Language typology and universals* 59:4, pp. 317-349.

Savatkova, A. (2002). *Gornoye narechiye marijskogo yazyka.* [Hill dialect of Mari.] Berzsenyi Dániel Főiskola, Savariae.

Serebrennikov, B. (1960). *Kategorii vremeni i vida v finno-ugorskikh yazykakh permskoj i volzhskoj grupp.* [Tense and aspect categories in the languages of Permic and Volga groups.] USSR Academy of sciences publishing, Moscow.

Shcherbinin, I. (2016). Aspektual'no-vremennyye kharakteristiki glagola v gornomarijskom yazyke. [Tense and aspect characteristics of Hill Mari verb.] [Ms]. Lomonosov Moscow State University.

Sitchinava, D. (2009). Stremit'sya presekat' na kornyu: russkaja konstruktsija s bylo po korpusnym dannym. [The Russian bylo-construction in corpus data.] Kiselyova, K. (ed.), *Korpusnyje issledovanija po russkoj grammatike.* Probel, Moscow, pp. 362-396.

Sitchinava, D. (2013). *Tipologija plyuskvamperfekta. Slavyanskij plyuskvamperfekt.* [Typology of Pluperfect. Slavic Pluperfect.] AST-PRESS KNIGA, Moscow.

Squartini, M. (1999). On the semantics of the Pluperfect: evidence from Germanic and Romance. *Linguistic typology* 3:1, pp. 51-89.

Testelets, Y. & M. Halilov (2002). Vyrazheniye evidentsial'nosti v bezhtinskom yazyke. [Expression of evidentiality in Bezhta.] Paper presented at: XI tutorial of the European society of Caucasian scientists, Moscow, June.

Tsypanov, E. (2005). *Grammaticheskiye kategorii glagola v komi yazyke.* [Verbal grammatical categories in Komi]. Institut yazyka, literatury i istorii komi, Syktyvkar.

Voronov, M. & D. Mordashova (2016). Gornomarijskiye glagol'nyje konstruktsii s ə̂l'ə̂ / ə̂lə̂n v zone sverkhproshlogo i za jejo predelami. [Hill Mari constructions with ə̂l'ə̂ / ə̂lə̂n within the realm of discontinuous past.] Paper presented at: XIII conference in typology and grammar for young researchers, Saint-Petersburg, November.
http://youngconfspb.com/application/files/9114/8055/4295/Mordashova_Voronov.pdf

Causing resultative passives

Michael Wilson

Originally, resultative passives were argued to lack agentivity (Embick 2004). Recent work argues the opposite (Meltzer-Asscher 2012; McIntyre 2013; Alexiadou et al. 2014, 2015; Anagnostopoulou 2003, 2013; Bruening 2014; Gehrke 2015). I argue these inconsistencies are best explained by treating the resultative passive as causative rather than agentive, an analysis which fits well with existing analyses of resultative structures (Kratzer 2005; Folli & Harley 2005; Schäfer 2012b). This account also offers a principled explanation for why certain unaccusative verbs but not others can form resultative passives (Pesetsky 1995; Levin & Rappaport Hovav 1995; McKoon & Macfarland 2000; Maienborn 2009; Gese et al. 2011).

1. Introduction

The passive is a syntactic construction in which (usually) the object of a verb appears in the canonical subject position, and the subject either appears as an oblique argument or not at all. Figuring out the syntactic restrictions on and the semantic consequences of forming passives has led to a typology of passives. Wasow (1977) began this line of inquiry by dividing passives into verbal and adjectival varieties. These usually share the same form, but differ in their meanings:

(1)   The glass was broken.
  a. *Verbal Passive*
       Someone broke the glass.
  b. *Adjectival Passive*
       The glass was in a broken state.

More recently, Kratzer (2000) and Embick (2004), among others, have argued for further dividing the class of adjectival passives. Kratzer (2000) divides adjectival passives into target state and resultant state passives based on whether the state described is reversible (e.g., *hidden*) or not (e.g., *proven*). Embick (2004) divides adjectival passives into statives and resultative passives based on whether the participle describes merely a state, or a state resulting from an event:

(2)   The door was closed.
     a.   *Eventive (= Verbal) Passive*
        Someone closed the door.
     b.   *Resultative Passive*
        The door was in a state of having become closed.
     c.   *Stative*
        The door was in a closed state.

I am concerned with Embick's (2004) classification. Embick's statives are constructed by attaching an adjectivizing head to a root. Resultative passives, on the other hand, are formed by attaching an aspectual head to a verbal projection formed from the root. In the case of resultative passives, the verbalizing head is non-agentive, leading to a passive that encodes eventivity sans agentivity. Finally, verbal passives are formed from a structure that includes v_{AGENT}, leading to eventivity and agentivity.

Recently, many have disputed Embick's claim that the resultative passive is non-agentive (Meltzer-Asscher 2012; McIntyre 2013; Alexiadou et al. 2014, 2015; Anagnostopoulou 2003, 2013; Bruening 2014; Gehrke 2015). They show evidence that Embick's diagnostics for (lack of) agentivity are incorrect, or that his data are insufficiently broad. Yet it *is* the case that diagnostics for agentivity are restricted with resultative passives compared to eventive passives. The availability of elements diagnosing agentivity requires either that the agent be responsible for the continuation of the state, or that the sortal features of the agent be crucial to the nature of the state (McIntyre 2013; Gehrke 2015; Alexiadou et al. 2015). These restrictions are usually attributed to the semantics of the aspectual head responsible for creating the resultative passive.

In this paper, I argue that these restrictions can be unified in analysis treating the resultative passive as causative. That is, the resultative passive is formed from a structure containing a v^0 introducing a causing event argument (cf. Pylkkänen 2002, 2008). I argue that this explains the restrictions on agent-related phrases with resultative passives, as well as the possible readings they can receive. Furthermore, it provides evidence that, contra Pylkkänen (2002, 2008), v_{CAUSE} and v_{AGENT} should be treated as separate heads in English (Alexiadou 2014; Folli & Harley 2005; Alexiadou & Schäfer 2006; Schäfer 2012b), in that this would account syntactically for the presence of resultative passives with unaccusative internally-caused change-of-state verbs (henceforth CoS verbs) (Levin & Rappaport Hovav 1995; McKoon & Macfarland 2000; Maienborn 2009; Gese et al. 2011). These verbs lack an agent argument, yet they can productively form resultative passives. The key, I argue, is the syntactic presence of a v_{CAUSE} head, which the resultative passive head attaches to a projection of.

I begin in §2 by reviewing the distinction between statives, resultative passives, and eventive passives. In §3, I review Embick's (2004) evidence for a lack of agentivity in the resultative passive in contrast with more recent evidence in favor of it (Meltzer-Asscher 2012; McIntyre 2013; Alexiadou et al. 2014, 2015; Anagnostopoulou 2013; Bruening 2014; Gehrke 2015). In §4, I argue that an account relying on causativity fares better in explaining these data, and relate my proposal to previous work on the licensing of v_{CAUSE} (Schäfer 2012b; Folli & Harley 2005). In §5, I discuss a prediction this account makes about when resultative passives should be possible, and show that this is borne out by examining which unaccusative verbs can productively form resultative passives in English. In §6, I briefly conclude.

2.   *Resultative passives*

2.1.   *Distinguishing resultatives and statives*

Resultative passives are distinct from eventive passives and statives because they encode both an event and a state. Evidence that resultative passives encode events is illustrated in (3) by contrasting *opened*, which has a resultative reading, and *open*, which is its stative counterpart:

(3)   a.   The package remained carefully opened/*open.
        b.   This door was built *opened/open.
        c.   John kicked the door *opened/open.
        d.   The door looks unopened/*unopen.

First, (3a) shows that resultatives, but not statives, can occur with adverbs. This shows that they encode events. The stative lacks an event, and correspondingly disallows adverbs.

Second, the stative can follow a verb of creation (*build*, *create*, *make*), while the resultative cannot, shown in (3b). The resultative *opened* cannot follow a verb of creation because it expresses a state resulting from a previous event. This is incompatible with a verb of creation, since the door did not exist prior to its being built, and so could not have undergone a previous event. In contrast, the stative expresses no prior opening event, and so can occur with verbs of creation.

Third, statives can appear as resultative secondary predicates, while resultatives cannot, shown in (3c). This is because the resultative encodes the state the door was in as being the result of an opening event, while here it is the result of a kicking event. In other words, one cannot kick the door into a state resulting from a prior event of its opening, only into a open state resulting from a kicking event.

Fourth, resultatives can productively take on negative *un-* prefixes, shown in (3d). With statives, this is only possible for particular lexical items (e.g., *unnecessary*, *unhappy*, but **unred*, **unsad*). This appears *prima facie* unrelated to the fact that resultatives encode an event, but it provides another way of distinguishing resultatives from statives.

2.2.   *Distinguishing resultatives from eventives*

In addition to an event, resultatives also encode a state, as I will now show. This establishes the resultative passive as distinct from both the stative (since it encodes an event) and the eventive passive (since it encodes a state). Since the resultative passive and the eventive passive always share the same form (Embick 2004), we must rely on evidence from the possible readings available in the sentences below.

(4)   a.   The door { remained / seemed } opened.
        b.   The door is being opened.
        c.   The door is (*being) unopened.

Sentence (4a) can only receive a reading where it describes a state that the door is in. In

other words, it cannot mean something like 'someone seemed / remained to open the door.' This shows that resultatives can occur with copular *remain* and with *seem* — like adjectives — while eventives cannot (Embick 2004; Sleeman 2014).

Sentence (4b) only has an eventive reading available; it can only mean 'someone is opening the door', not 'the door is being in a state of being opened'. In other words, resultative readings are unavailable in the progressive aspect, because states cannot be progressive.

Finally, shown in (4c), *un-* prefixation is unproductive with eventive passives (Sleeman 2014), though possible again for particular lexical items (*being unzipped*, *being undone*, but *being unread, *being unused). Further, we can note that in cases when *un-* prefixation is possible with eventive passives, it always carries a reversative meaning as it does with verbs, but for resultative passives it always carries a negation meaning, as it does with adjectives.

3.   *Agentive resultatives*

To this point, I have reviewed evidence that the resultative passive encodes both a state and an event, distinguishing it from both statives and eventives. Embick (2004) claims that the resultative passive is further distinguished from the eventive passive in that it has no implicit agent.

(5)   a.   The metal is hammered by John.
        b.   The ship is sunk to collect the insurance.

Only an eventive reading is available for (5a), which is habitual in the present tense. That is, it cannot mean 'the metal is in a state of having become hammered by John', according to Embick (2004). The *by*-phrase can only be present if there is an implicit agent, as it identifies the implicit agent with its referent. Since a resultative reading is impossible with a *by*-phrase, we must assume that there is no implicit agent in such cases.

Sentence (5b) also only has an eventive reading, forced by the presence of a purpose clause that the implicit agent controls into (cf. Roeper 1987; Baker et al. 1989). It cannot mean 'the ship is in a state of having been sunk so that the sinker could collect the insurance.' Embick (2004) argues that while eventive passives have a syntactically present implicit agent that can function as a controller, resultative passives lack this argument. If there is no implicit agent, PRO cannot be bound, explaining why a resultative reading with a purpose clause containing it is unavailable.

Third, unlike eventives, resultatives allow for reflexive interpretations (Baker et al. 1989; Kratzer 2002; Bruening 2014):

(6)   a.   The children are washed. They did it all by themselves.
        b.   The children are being washed. #They did it all by themselves.

In (6a), a continuation specifying coreference between the agent and theme of the washing event is possible under a resultative reading of the participle. In contrast, such a reading is not possible in (6b); with an eventive passive (forced by the progressive aspect), the washer cannot be the wash-ee. Again, this seems to show that the eventive passive has a syntactically present implicit external argument which cannot be bound by *the children*, resulting in a disjoint reference effect.

The resultative seemingly has no implicit agent, so coreference is not ruled out.

Recently, however, the data motivating these claims have been called into question. In particular, resultatives do seem to be able to combine with agent-oriented modifiers and *by*-phrases in some contexts. Similarly, disjoint reference effects do not occur for some verbs — implying that an implicit agent may be present in such cases.

For example, McIntyre (2013) notes that while *by*-phrases are degraded if added to sentences like (4a), they are acceptable if the referent they introduce is 'responsible for continuing the state expressed by the participle' (7) or if its sortal features are crucial to the nature of the state (8) (examples and judgments from McIntyre 2013).

(7)   a.   The road remained { blocked by police / supported by pylons }.
       b.   The dictator remained { unsupported / propped up / underestimated } by the warlords.
       c.   Edeltraud seemed flattered by { the report / ??the journalist[1] }.

(8)   a. %The text seems written by a { genius / foreigner / ghostwriter }.
       b.   *The text seems written by John. He wouldn't have gone home without finishing it.

Bruening (2014) shows that *by*-phrases are particularly productive with *un-* prefixed resultatives; these contexts may highlight the relevance of the sortal features of the referent:

(9)   a.   Little nature remains untouched by human hands.
       b.   (Invading Commander:) I want the treasury left untouched by anyone but me!
       c.   Toddler unfazed by lion encounter. (headline)

Resultative passives can also appear with instrumental *with*-phrases, which are taken to show the presence of an agentive/animate external argument (Bruening 2014; Alexiadou et al. 2015).

(10)   a.   The radioactive nucleotides are so small that they remain unseen, even with the most powerful microscope.
       b.   ...the very earliest stages of an arteriosclerotic plague, which remains undiscovered even with the most up to date clinical diagnostics.
       c.   ...the round edge looks beaten with a hammer or sledge...
       d.   ...the stone looks cut with some kind of machine...
       e.   ...interior still looks carved with a dull spatula...

In some cases, control into purposes clauses is possible with resultatives (McIntyre 2013), which could be taken to show that an implicit agent is syntactically present (but see Williams 1974 for an opposing view).

(11)   a.   The ... bags remained closed in order to keep the modified atmosphere intact.
       b.   Use of the name Blohm + Voss remained prohibited, in order to ... spare the world the shock that ships were being built there again.
       c.   The investigation ... remained limited in order to protect the police.

---

[1]Troy Messick (p.c.) reports to me that he accepts *the journalist* here, but possibly only with coercion, such that *the journalist* stands for *what the journalist said*.

A web search revealed that agent-oriented adverbs can occur with resultatives.[2]

(12)   a.   ...the new "Common Core" curriculum..seems intentionally designed to make...[3]
        b.   Ms. Morgan seems willingly buried within the songs.[4]
        c.   ...there's a two inch long section of wire that seems intentionally un-insulated![5]

Finally, McIntyre (2013) shows the following examples of resultatives that strongly disfavor[6] a reflexive interpretation.

(13)   a.   #John criticized himself, but to me he seemed unfairly criticized.
        b.   #Some people trust themselves while others underrate themselves and think they won't succeed. Mary seems very underrated and not very trusted.
        c.   #He had self-hate problems and remained very hated until he sought help.

Bruening (2014) and McIntyre (2013) both conclude that the (un-)availability of coreference in these cases is related to particular verb classes, rather than to a particular structure/reading. The relevant verb classes are those identified in Schäfer (2012a): naturally and inherently reflexive verbs. Naturally reflexive verbs are those describing events where the agent is typically coreferential with the theme, such as *wash*, *shave*, etc. Inherently reflexive verbs are grammatically reflexive; such verbs do not occur in English, but can be found in other languages (e.g., Spanish *reirse* 'laugh-REFL'). Naturally and inherently reflexive verbs allow coreferential interpretations in resultative passives, while naturally disjoint verbs do not. I will later relate this fact to my proposal that resultative passives involve v_{CAUSE}.

To sum up, a series of diagnostics shows that resultative passives, contra Embick (2004), do display characteristics associated with agentivity, at least in some contexts. *By*-phrases are possible when the referent is responsible for continuing the state the participle expresses or when the referent's sortal features are crucial to the nature of the state, and are especially productive with *un-* prefixed resultatives. Instrumental *with*-phrases and agent-oriented adverbs can also occur with resultative passives. This casts doubt on Embick's (2004) analysis that the resultative passive is non-agentive.

---

[2] These examples sound perfectly natural to me, as well as to other native English speakers to whom I have presented these data.

[3] *http://www.naturalnews.com/044338_Common_Core_math_education_Americas_children.html*

[4] *http://www.stephaniesid.com/press/*

[5] *http://www.ford-trucks.com/forums/1370192-bare-wire-in-distributor-harness.html*

[6] McIntyre claims that reflexive readings are impossible in these examples. I do not share this judgement, nor do others to whom I have presented these sentences. The reflexive readings are available for me. Troy Messick (p.c.) notes that (i) sounds good to him:

(i)   John is unfairly criticized, but he mostly does it to himself.

This supports my argument that coreference is available in these kinds of sentences, but simply strongly dispreferred for reasons relating to world knowledge.

4.   Causative resultatives

We now seem to be left with a puzzle: in some contexts, elements diagnosing agentivity are incompatible with resultative passive readings, but in other contexts, they are not. In this section, I provide evidence that resultatives are causative rather than agentive, by examining the readings of the above sentences more closely. To some extent, this explains the different results of the diagnostics above.

4.1.   Readings of the resultative passive with transitive verbs

Let us return to McIntyre's (2013) and Bruening's (2014) examples that aim to show an implicit agent is present in resultative passives. Considered more carefully, these support a causative rather than an agentive reading. I begin with McIntyre's (2013) examples, repeated below.

(14)   *By-phrase Referent Responsible for Continuing the State*
    a.   The road remained { blocked by police / supported by pylons }.
    b.   The dictator remained { unsupported / propped up / underestimated } by the war-lords.
    c.   Edeltraud seemed flattered by { the report / ??the journalist }.

(15)   *Sortal Specification Responsible for the Nature of the State*
    %The text seems written by a { genius / foreigner / ghostwriter }.

Consider example (14a) in detail. In a transitive use, *blocked* has two readings; one only allows agent subjects, and one allows causer subjects. These are illustrated in (16).

(16)   a.   *Agentive blocked*
        The police blocked the road (with traffic cones, and then left for the day).
    b.   *Causative blocked*
        i.   Fallen trees blocked the road.
        ii.   The police blocked the road (for hours).

The causative reading of *blocked* is only compatible with a scenario when the causer is present during the blocking event. Similarly, the resultative passive is only possible in such a case:

(17)   a.   *Agentive*
        The police blocked the road with traffic cones, and left for the day. Even though they were supposed to, they forgot to come pick them up in the afternoon.
        #In fact, that evening, the road remained blocked by police.
    b.   *Causative*
        The police blocked the road at dawn, and it seemed like they would never leave. In fact, that evening, the road remained blocked by police.

A similar analysis explains (14b): *support* and *prop up* both show the same agentive-causative ambiguity as *block*. The exception in this group is *underestimate*, which resists a

non-animate subject. This could lead us to conflate the *by*-phrase referent with an agent. We should resist this temptation. For one thing, it is unclear that *underestimate*'s external argument should be considered an agent, as opposed to an experiencer (Rajesh Bhatt, p.c.). For another thing, we may attribute restrictions on the kind of entity that can serve as an external argument to *underestimate* to encyclopedic knowledge rather than syntax; given its meaning, *underestimate* requires an animate subject, which is easily conflated with an agent. We can see similar restrictions based on encyclopedic knowledge with periphrastic causatives, which are supposed to be indicative of true causers (Levin 2009):

(18)   a.   Mary made John hire a lawyer.
        b. ??The wind made John hire a lawyer.

Presumably, what we know about events of underestimating and making someone hire a lawyer is that they always or typically involve an animate causer. Animate causers are also indicative of agentivity, explaining why we might confuse these with agents, though they are not. My proposal is thus along the lines of Schäfer's (2012b) analysis of oblique causers, which must be human but are not agents. The fact that such causers must be human derives from the fact that they are introduced by an applicative head that expresses a possessive relation, which can only hold of humans. Nothing different is going on in the case of the resultative passive.

McIntyre's (2013) reported judgements of (14c) also corroborate the idea that the resultative passive is causative, since we see that the sentence is perfectly acceptable with a causer (*the report*), but somewhat strange with an agentive causer (*the journalist*). This follows if we assume that allowing an agentive causer when encyclopedic knowledge favors a non-agentive causer requires some sort of coercion. In cases where encyclopedic knowledge only permits agentive causers (or at least strongly favors them), this cost is diminished.

Adopting an analysis where the encyclopedic knowledge of the verb plays a role in determining what can serve as a possible causer also allows us to explain the variability in disjoint reference effects with different verbs. As mentioned before, those verbs that are inherently and naturally reflexive (Schäfer 2012a) allow coreferential readings in the resultative passive, while those that are naturally disjoint strongly disfavor them.

(19)   a.   The children seem washed. They did it themselves.
        b.   The children seem hated. ??In fact, they hate themselves.

If these involve an implicit causer whose referent is determined by encyclopedic knowledge, this contrast is explained. Verbs that describe events where the theme is typically coreferential with the agent/causer allow readings where the implicit causer argument is understood as coreferential with the theme easily. Verbs that describe events where the theme is typically non-coreferential with the agent/causer disprefer such readings. Note that McIntyre's (2013) examples of disjoint reference effects with resultative passives in (13) all involve naturally disjoint verbs: *criticize*, *underrate*, *trust*, and *hate*.

Regarding the availability of a resultative passive reading when the sortal specifications of the *by*-phrase referent is crucial to the nature of the state, as in (15), I suggest that it is because only at this level of granularity can we consider its referent a causer — what is relevant is not the specific referent's identity, but rather the fact that the referent is the direct cause of a result

state. Consider the following examples (Rajesh Bhatt, p.c.):

(20)   *I know there were cookies in the break room, but when I walk in I find they are gone. I also smell a cologne that I know only John wears.*

    a.   *Eventive passive*
        The cookies seem to have been eaten by John.
    b.   *Resultative passive*
        #The cookies seem eaten by John.

(21)   *I know there were cookies in the break room, but when I walk in I find they are half-eaten. I notice that pattern of bites in what remains of the cookies is consistent with John's having eaten them, because he has no teeth.*

    a.   *Eventive passive*
        The cookies seem to have been eaten by John.
    b.   *Resultative passive*
        The cookies seem eaten by John.

What is relevant is that we must be able to 'read off' John's effect on the cookies from the resulting state they are in itself to license a resultative passive. No such restriction holds for the eventive passive.

Borrowing a notion of direct cause from Kratzer (2005) might allow for an explanation of these restrictions: an eventuality directly causes another just in case there is no eventuality intervening between them. Put another way, the end of one eventuality is the beginning of another. Here, the end of the eating eventuality John participated in is the beginning of the state of the cookies having been eaten if John is the direct cause. But since John is a participant in the eating eventuality, and the end of the eating eventuality is the beginning of the state holding of the cookies, John is also present in the state holding of the cookies. In order for us to consider John a direct cause of the state holding of the cookies, we must be able to infer his role from the state. This explains why we can only use sentences like (20b/21b) in cases when the sortal specifications of the referent are crucial to the nature of the state. In such cases, the referent is present in the state in some way. In cases when the *by*-referent is responsible for continuing the state (e.g., 14), it is trivially true that the referent of the *by*-phrase is present in the result state, which characterizes causative uses of verbs like *block*. Analyzing the resultative passive as causative, then, gives us a unified account of McIntyre's (2013) generalizations about when *by*-phrases are possible.

McIntyre (2013) himself advocates for a semantics where the passive participle morpheme introduces a CAUSE relation, and Meltzer-Asscher (2012) also includes a CAUSE relation for resultative passives (though it is not part of the participle morpheme). The difference between my proposal and theirs is that I relate the restrictions on *by*-phrases to the nature of the (direct) CAUSE relation in combination with world knowledge and lexical semantics directly, rather than to other factors. My approach is in this respect the same as García-Pardo's (To appear), who independently develops an analysis in which only causative verbs may form adjectival (resultative) passives, and in which the notion of causativity plays a central role in determining the availability of allegedly agent-oriented modifiers in resultative passives. I unfortunately lack

the space here to examine the fine-grained differences between my approach and García-Pardo's (To appear) in the detail they merit, though I will examine them in a follow-up paper.[7]

However, the kind of approach that I and García-Pardo (To appear) take and the kind of approach that Meltzer-Asscher (2012) and McIntyre (2013) take differ in that the former kind explicitly prohibits agents from occurring with resultative passives. Meltzer-Asscher (2012), in contrast, assumes an implicit agent in resultative passives along with the CAUSE relation. Similarly, McIntyre (2013) argues for the presence of an implicit initiator, ranging over agents and causers, in addition to the CAUSE relation. I, like García-Pardo (To appear), instead support an approach where only causers and not agents occur with resultative passives, since we can only have causative readings of resultative passives formed from verbs that are ambiguous between agentive and causative uses. Encyclopedic knowledge will sometimes lead to cases where it is hard to distinguish between the two on the basis of features like animacy or intentionality, as in the case of *underestimate*, but this does not mean that agents and causers are the same.

4.2.   *Theoretical motivations for a causative analysis*

In addition to the empirical generalizations discussed in the preceding section, we can motivate a causative analysis of the resultative passive by noting that typical analyses of (non-passive) resultatives treat them as causative (e.g., Kratzer 2005 *inter multa alia*). Adopting a causative analysis of the resultative passive allows us to straightforwardly claim that the resultative passive is simply the passive of the resultative. Folli & Harley (2005) and Schäfer (2012b) further link resultative structures and causativity, showing that causers are licensed only in resultative contexts. For example, Folli & Harley note that canonical nominative causers can only occur in resultative contexts. Such contexts either involve a verb that lexically specifies a result state, or a structure with a resultative secondary predicate (examples from Folli & Harley 2005).

(22)   *Lexically Causative*
    a.   The sea destroyed the beach.
    b.   The groom destroyed the wedding cake.

(23)   *No Resultative Secondary Predicate, Non-Lexically Causative*
    a.   *The sea ate the beach.
    b.   The groom ate the wedding cake.

(24)   *Resultative Secondary Predicate, Non-Lexically Causative*
   The sea ate away the beach.

Agents and causers may occur with lexically causative verbs, which express a resultative event structure. Verbs of consumption (here, *eat*), are not resultative, and license agents but not causers. If a verb of consumption is combined with a resultative secondary predicate, forming a resultative event structure, then causers are licit, as in (24). Furthermore, Folli & Harley allow for causers to be either [+HUMAN] or [−HUMAN], as I argued for in the preceding section.

---

[7]García-Pardo (2017) gives a rather more different, Aktionsart-based account of adjectival (resultative) passives, using Ramchand's (2008) framework. I also lack the space to discuss how this analysis differs from mine.

Schäfer (2012b) notes that oblique causers and PP causers are subject to the same restriction: they can only occur in the presence of a resultative event structure.[8]

Syntactically, Folli & Harley's (2005) analysis is realized by having v_{CAUSE} c-select a small clause complement. The lexical verb can be analyzed either as incorporating into $v_{\text{CAUSE}}/v_{\text{DO}}$ or else as projecting a process vP below them — the exact mechanism is unimportant.
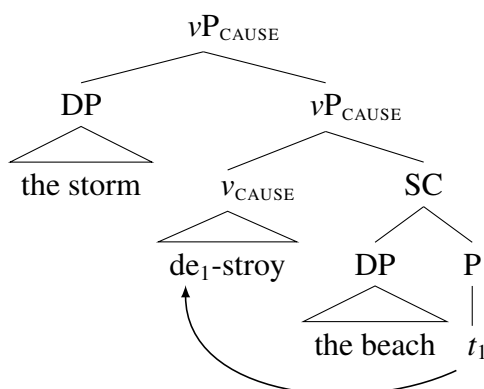
(25) a. *Agentive v*P
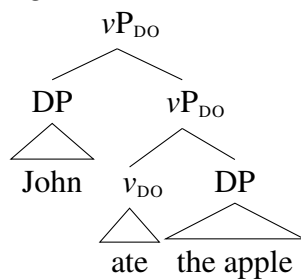
  b. *Causative v*P
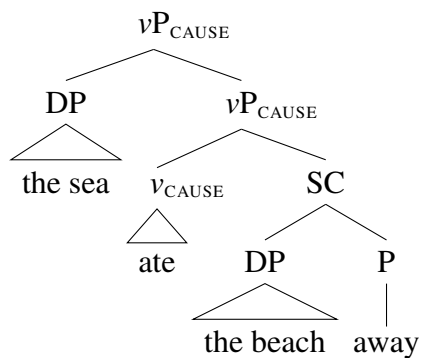
Folli & Harley propose that lexically causative verbs underlyingly involve a resultative event structure, wherein a bound morpheme incorporates into the root verb. For example, the underlying structure for the lexically causative verb *destroy* is the following:

(26)

[8]For Schäfer, the syntax of each construction is slightly different: canonical nominative causers are introduced by a Voice[0] taking v_{CAUSE} as a complement, oblique causers are introduced by an applicative head with a v_{CAUSE} complement, and PP causers are introduced by a PP in the specifier of v_{CAUSE}. These details are not relevant to my analysis since oblique causers do not occur in English, and PP causers are not necessarily diagnostic of syntactic causativity in light of Levin's (2009) data regarding *from*-phrases.

In light of this finding, we must note that, as expected, resultative passives involve a resultative eventive structure. Minimally, there must be some event that occurred, and a result state following it. In cases when there is no lexically encoded result state, one must be coerced in a resultative reading. For example, Embick (2004) notes that out of context, a resultative reading of (27) is odd.

(27)   The tires are kicked.

However, 'given a scenario in which I work in a tire factory, and I have to kick all of the tires before I can go home, it becomes much better' (Embick 2004:361). That is, in a scenario where 'being kicked' is the relevant result state of a kicking event, the resultative reading is available. More generally, a resultative passive requires a resultative event structure, which is linked to causativity.

4.3.   A causative semantics for the resultative passive

Given the close tie between resultative structures and causativity, it is unsurprising that the resultative passive would be causative. Such a structure might look something like this, following Folli & Harley (2005).

(28)

```
                        AspP
                       /    \
                  Asp_R      vP_CAUSE
                            /        \
                      v_CAUSE         SC
                      /     \        /   \
               v_CAUSE  ROOT     DP      X
                          |     /  \     |
                       √KICK  the tires  R
```
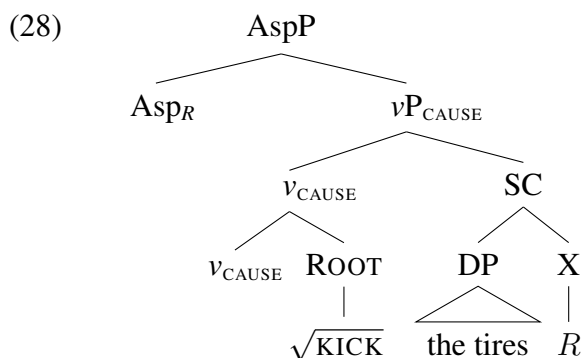
where R is a free result state variable that is interpreted pragmatically[9] and Asp_R existentially binds a causing event:

(29)   a. $[\![R]\!] = \lambda x.\lambda e.R_i(x, e)$
         b. $[\![v_{\text{CAUSE}} \text{ ROOT}]\!]^{10} = \lambda f_{\langle s, t \rangle}.\lambda e.\exists e':f(e') \wedge \text{CAUSE}(e, e') \wedge \text{kick}(e)$
         c. $[\![\text{Asp}_R]\!] = \lambda f_{\langle s, t \rangle}.\exists e:f(e)$

Putting this together gives us the following denotation for *the tires are kicked*:

(30) $\exists e,e':R_i(\textit{the.tires}', e') \wedge \text{CAUSE}(e, e') \wedge \text{kick}(e)$

This asserts that there is an eventuality *e* and an eventuality *e'* such that a pragmatically given

---

[9]Cf. the free relation variable proposed for an analysis of free readings of the English *-'s* in Partee (1983/1997), as well as Maienborn's (2009) and Gese et al.'s (2011) analyses of the resultative passive.
[10]Cf. Pylkkänen (2002, 2008).

relation holds between the tires and e', e directly causes e', and e is a kicked event. In other words, a pragmatically determined state holds of the tires that was directly caused by a kicking event. In such a context, the most natural reading is one where the state is a state of having been kicked. If we further specify the result state, we can get resultative passives like *the tires are kicked flat*, *the tires are kicked over*, and so on. By 'directly causes' I mean Kratzer's (2005) notion, as described earlier: no eventuality may intervene between the kicking eventuality and the state holding of the tires.

Other cases will involve a more typical result state in place of R, like *broken*:

(31) $[\![\,broken\,]\!] = \lambda \mathrm{x}.\lambda \mathrm{e}.\mathrm{broken(x, e)}$

Putting this together with the denotations above yields the following for *the glass is broken*:

(32) $[\![\,the\ glass\ is\ broken\,]\!] = \exists \mathrm{e,e'}{:}\mathrm{broken}(\mathit{the.glass'}, \mathrm{e'}) \wedge \mathrm{CAUSE(e, e')}$

This asserts that some event caused the glass to enter a broken state.

By-phrases, agent-oriented adverbs, and instrumental phrases all modify the causing event e below Asp_R. Such modifiers are predicted to be compatible only in cases when encyclopedic knowledge requires or favors a causer that can also be an agent, as argued for above.

5.   *Resultative passives of unaccusatives*

An important difference between my proposal and McIntyre's (2013) is that he introduces the notion of causativity in the semantics of the passive participle itself, while I argue that the CAUSE relation is introduced by v_{CAUSE}, which is separate from the passive aspectual head. Then suppose, as I have argued, Asp_R from the preceding section c-selects for a projection of v_{CAUSE}, and cannot occur with v_{AGENT}. If this is the case, we make the following prediction: resultative passives should be allowed with all verbs that can be shown to independently encode causativity but not agentivity. Are there such verbs?

I assert that unaccusative internally-caused change-of-state verbs (CoS verbs) are such verbs (Levin & Rappaport Hovav 1995; McKoon & Macfarland 2000). McIntyre (2013), Pesetsky (1995:116ff), and Gese et al. (2011) have noted that passives can occur with some but not all unaccusative verbs. McIntyre (2013) argues that the relevant class of verbs are those that designate a 'salient, relatively stable' result state, while Gese et al. (2011) argue that pragmatics is the relevant factor, at least in German. While I cannot speak to the German data, since there is evidence that the structure of resultative passives can vary considerably across languages (see especially the discussion of Greek in Alexiadou et al. 2015: ch. 5), I will address McIntyre's (2013) argument by showing that the presence of v_{CAUSE} is a better predictor. This follows Pesetsky's (1995) suggestion that those unaccusative verbs allowing passive uses are just those that we can conceive of as projecting an ambient/internal causer argument — which, under my analysis, should allow them to form resultative passives. I refine Pesetsky's idea by showing that it is not enough that we can conceive of an ambient/internal cause of the event, but that the verb must be syntactically causative, which can be shown by independent diagnostics.

5.1. *Diagnostics for causativity and CoS verbs*

One proposed diagnostic for syntactic causativity for unaccusative verbs is the ability to take a *by itself* phrase meaning 'without any particular cause' (Koontz-Garboden 2009; Alexiadou et al. 2015). This distinguishes unaccusative uses of verbs from (eventive) passive uses:

(33)   a.   The ship sank (by itself).
        b.   The ship was sunk (*by itself).

However, this diagnostic must be applied to resultative passives carefully, as we certainly cannot say *the ship remains sunk by itself*. The semantics for Asp_R given in §4.3 might provide an explanation of why *by itself* fails to diagnose causativity in the case of resultative passives, even though it does so for unaccusative verbs.

As Alexiadou et al. (2015) and Schäfer (2007) note, the use of *by itself* indicates that the speaker cannot identify anyone or anything that caused the antecedent to undergo the event. Furthermore, it is also the case that the antecedent did undergo the event. That is, a sentence with *by itself* involves both asserting that an event occurred, while denying the presence of a causer of that event. Following this line of reasoning, suppose that *by itself* attaches to a projection of v_{CAUSE} above its SC complement and has the following semantics:

(34) $[\![$ *by itself* $]\!] = \lambda f_{\langle s, t \rangle}.\exists e{:}f(e) \wedge \neg\exists x{:} x$ is a participant of e

Combined with the denotation of v_{CAUSE} above and a SC like [*the glass broken*], this would yield the following:

(35)   a. $[\![$ $v\text{P}_{CAUSE}$ *the glass broken* $]\!] = \lambda e.\exists e'{:}broken(\textit{the.glass}', e') \wedge \text{CAUSE}(e, e')$
        b. $[\![$ $v\text{P}_{CAUSE}$ *the glass broken by itself* $]\!]$
            $= \exists e,e'{:}broken(\textit{the.glass}', e') \wedge \text{CAUSE}(e, e') \wedge\neg\exists x{:} x$ is a participant of e

More idiomatically, this asserts that there is an eventuality *e* that is the cause of an eventuality *e'* of the glass being broken, and there is no entity *x* that participated in the causing event *e* (cf. Schäfer 2007).

Adding Asp_R to this would create a type mismatch — it requires a function of type $\langle s, t \rangle$, but here we only have something of type *t*. This would account for why *by itself* cannot be used with the resultative passive, despite the resultative passive being causative. Nevertheless, we might see whether a verb can occur with *by itself* in a non-resultative-passive structure, and use that to determine whether it is causative. If it is, then we predict that it should be able to form a resultative passive.

Deal (2009) argues for another diagnostic for causativity: unaccusative verbs that permit expletive *there* are non-causative; those that do not permit it are causative. This is corroborated by the fact that unaccusative verbs that allow *there*-insertion do not allow the 'with no particular cause' reading of *by itself*:

(36)   a.   The ice cream melted by itself.
        b.   *There melted some ice cream in the heat.

(37)   a.   #The portrait hung on the wall by itself. (Okay under an 'alone' reading.)
        b.   There hung a portrait on the wall.

So we have established two diagnostics for syntactic causativity for unaccusative verbs: the availability of a *by itself*-phrase (at least in a non-passive context), and the lack of availability of *there*-insertion.[11,12] Let us apply these to two unaccusative verbs: *break* and *arrive*.

(38)   a.   The vase broke by itself.
        b.   *There broke a vase on the table.

(39)   a.   #The letter arrived by itself. (Okay under an 'alone' reading.)
        b.   There arrived a letter in the mail.

According to these diagnostics, then, *break* is causative, and so should be able to occur in the resultative passive. In contrast, *arrive* is not causative, and so should not be able to occur with the resultative passive. This prediction is borne out for these and other verbs:

(40)   a.   The vase seems (un)broken.
        b.   *The letter seems (un)arrived at the house.
        c.   *The board seems (un)stood against the wall.
        d.   *The wizard seems (un)appeared in the castle.
        e.   *The boy seems (un)run to the store.

Of course, there is a possible confound here: *break* can alternate between a transitive use and an unaccusative use, while the other verbs here only occur intransitively. If the resultative passive is formed from the transitive use, we have only shown that resultative passives can occur with transitive verbs but not intransitive verbs.

However, there are some non-alternating unaccusative verbs that pass the diagnostics for causativity: these are the CoS verbs (Levin & Rappaport Hovav 1995).

(41)   a.   The log decayed by itself.
        b.   *There decayed a log in the forest.
        c.   *The humid weather decayed the log.

---

[11]A problem with these diagnostics is that they cannot be applied to transitive verbs, which is why I did not introduce them above.

(i)   a.   #John kicked the tires by himself. (Okay under an 'alone' reading.)
       b.   *There kicked John the tires.

Both *there*-insertion and *by itself* are out, which leads to a contradiction: the unavailability of *by itself* suggests that the verb is non-causative, while the unavailability of *there*-insertion suggests that it is. Presumably, these are in fact ruled out here for independent reasons, which is why my argument earlier relied on possible readings of resultative passives formed from transitive verbs, rather than this kind of syntactic evidence.

[12]Another putative diagnostic for causativity that has been explored in Alexiadou et al. (2006), Alexiadou et al. (2015), and Levin (2009) is the ability of a predicate to license a *from*-phrase with a causer referent. Unfortunately, this diagnostic is difficult to apply clearly, since the *from*-phrase itself is most likely introducing the notion of causativity (Levin 2009), meaning it cannot diagnose whether syntactic causativity is present elsewhere in the sentence. Nevertheless, there are still restrictions on when *from*-phrases can introduce causing events (Deal 2009). The complications involved in using this as a diagnostic mean that I will put it aside here.

(42) a. The situation deteriorated by itself.
   b. *There deteriorated a situation at the conference.
   c. *The behavior of the conference attendees deteriorated the situation.

(43) a. The meat rotted by itself.
   b. *There rotted some meat in the kitchen.
   c. *The unhygienic conditions rotted the meat.

These verbs can productively form resultative passives, despite not having transitive uses:

(44) a. The log seems (un)decayed.
   b. The situation seems (un)deteriorated.
   c. The meat seems (un)rotted.

Recall that the ability to occur with *seem/remain* and the productivity of *un-* prefixation distinguish resultative readings from eventive and stative readings.

 Furthermore, unaccusative verbs that are ambiguous between an eventive reading and a stative reading only permit resultative passives with the eventive reading. These verbs include *bloom*, *blossom*, *flower*, *sprout*, *grow*, etc. As expected, only the eventive/causative uses of these verbs pass the *by itself* and no-*there*-insertion tests.

(45) *grow*: *Stative*

   a. #The corn grew by itself in our garden last year.
   b. There grew some corn in our garden last year. (Deal 2009: ex. 19a)

(46) *grow*: *Eventive*

   a. The corn grew very slowly in Massachusetts by itself.
   b. *There grew some corn very slowly in Massachusetts. (Deal 2009: ex. 19b)

(47) *bloom*: *Stative*

   a. #The rosebush bloomed on the patio by itself.
   b. There bloomed a rosebush on the patio. (Deal 2009: ex. 20a)

(48) *bloom*: *Eventive*

   a. The rosebush bloomed very slowly on the patio by itself.
   b. *There bloomed a rosebush very slowly on the patio. (Deal 2009: ex. 20b)

According to the diagnostics for causativity, then, these are causative in only their eventive use, but not their stative use. Correspondingly, resultative passives are only available with an eventive interpretation.

(49) The roses seem bloomed. =

   a. The roses seem to have bloomed.
   b. *The roses seem to be blooming.

(50) The corn seems grown. =

   a. The corn seems to have grown.
   b. *The corn seems to be growing.

Note that a verb must pass the *by itself* and no-*there*-insertion tests to allow a resultative passive. It is not enough that we can conceive of the event as causative. Take the examples below, adapted from Pesetsky's (1995) (325m,326e):

(51)   a.   *The paint looks glistened.
        b.   The boat looks capsized.

Pesetsky suggests that (51b) is licensed in a scenario where 'some property of the boat causes it to capsize', which allows *capsize* to project an A(mbient)-Causer role that allows it to be passivized. However, this wouldn't explain why (51a) is out: in the most natural case, it would be something intrinsic to the paint that causes it to glisten, which should allow us to project an A-Causer role, and thus allow for passivization. However, if we apply the diagnostics above, we note that *capsize* passes the tests for causativity, and *glisten* fails:

(52)   *By itself*

        a.   #The paint glistened by itself.
        b.   The boat capsized by itself.

(53)   There-*Insertion*

        a.   There glistened some paint on the wall.
        b.   *There capsized a boat on the ocean.

It is the tests for syntactic causativity that correlate with the availability of a resultative passive, not the ability to conceive of an intrinsic cause.[13]

It is also interesting to note that several of McIntyre's (2013) examples of resultative participles formed from unaccusative verbs include something like a resultative secondary predicate when they do not clearly fall into the class of CoS verbs:

(54)   backslidden people, flown-away bird, run-out license, fallen-over/caved-in building

This is exactly when the present account predicts that such participles should be available, given that resultative secondary predicates license v_{CAUSE}. Most of McIntyre's other examples fall into the class of CoS verbs:

(55)   rested/lapsed/returned/vanished people, deteriorated/decayed/capsized/rusted boat, wilted/faded flower, unerupted volcano, expired/elapsed license, stuck window, hatched chicken, swollen/bloated hand, collapsed building

These allow *by itself* modification and disallow *there*-insertion, which supports the analysis that

---

[13]Some of Pesetsky's (1995) examples in his (325–326) seem more problematic: in particular, he reports that *newly arrived packages* is well-formed, which seems indisputable. Similarly indisputable is the fact that *arrive* fails the tests for causativity, which under my account predicts that it should not be able to form resultative passives, shown in (40b). I cannot address this concern in detail, but I will note that Pesetsky's examples all involve pre-nominal adjectival uses of unaccusative verbs, which may not always be the same thing as resultative passive uses. In particular, *newly arrived packages* seems to have a perfective reading (i.e., the packages that *have* recently arrived), rather than a resultative reading (i.e., the packages that *are* recently arrived). In addition, removing *newly* makes participles with *arrive* unacceptable (*arrived packages*) suggesting that something else is going on.

they are causative. The only exceptions in McIntyre's list are *failed/Dutch-descended/retired people* and *defected spies*. These clearly disallow *there*-insertion, but they are also not particularly good with *by itself* insertion. I attribute this to encyclopedic knowledge that tells us such events must have some cause; spies cannot just defect 'on their own' — they must decide to do so, and so on.

McIntyre's (2013) idea that this should be attributed to a relatively stable result state *per se* is cast into doubt by the fact that the participle need not describe a stable result state in any particular case, as shown by the ability to productively add 'partially', 'somewhat', 'nearly', 'almost', 'half-', or 'not quite' to any given example:

(56) a. The log seems partially decayed.
      b. The flower seems nearly wilted.
      c. The rock seems almost eroded.

While there is a still a stable result state that these participles describe, they are not describing a stable result state in these examples. If the restriction is pragmatic and not syntactic/semantic, it is unclear why the fact that they can describe a stable result state should matter here. If we do want to say that the restriction is syntactic/semantic, then the explanation for why these allow resultative passives is essentially the same as the present account: the lexically specified result state is the small clause predicate.

A final potential critique to address could come from McKoon & Macfarland's (2000) work, which shows that some CoS verbs actually alternate. For example, some speakers may allow transitive uses of *bloom*, *stagnate*, *erode*, and so on. Maybe, then, resultative passives of these verbs are formed from their transitive uses. However, McKoon & Macfarland find some CoS verbs that have a very low probability of occurring in transitive uses in their corpus: *bloom*, *blossom*, *decay*, *deteriorate*, *flower*, *germinate*, *rot*, *stagnate*, and *wilt*. These all have a probability of occurring in a transitive construction of < 0.1. These all pass the *by itself* and no-*there*-insertion tests, and can form resultative passives:

(57) a. The roses seem { (un)bloomed / (un)blossomed / (un)wilted }.
      b. The log seems { (un)decayed / (un)rotted }.
      c. His health seems (un)deteriorated.
      d. The bush seems (un)flowered.
      e. The seed seems (un)germinated.
      f. ...total energy consumption seems stagnated within the EU...[14]

Since these verbs rarely or never occur in transitive constructions, it seems unlikely that the resultative passive is derived from some marginal transitive use. Instead, it is likely they have the following structure:

---

[14]*https://books.google.com/books?isbn=1317128141*

(58)
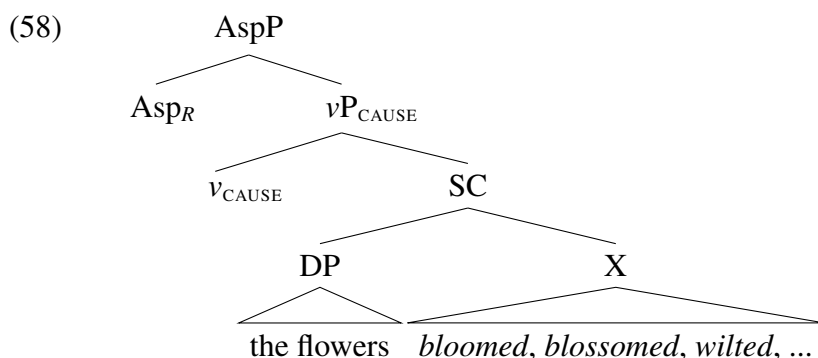
The semantics for such a structure is presented in (59), partially repeated from (29), with the semantics of unaccusative verb roots modeled after the semantics for *broken* in (31):

(59)   a.   { ⟦ *bloomed* ⟧, ⟦ *blossomed* ⟧, ⟦ *wilted* ⟧, ... }
            = { λx.λe.bloomed(x, e), λx.λe.blossomed(x, e), λx.λe.wilted(x, e), ... }
       b.   ⟦ v_{CAUSE} ⟧ = $\lambda f_{\langle s,\, t\rangle}.\lambda$e.$\exists$e′:f(e′) $\wedge$ CAUSE(e, e′)
       c.   ⟦ Asp_R ⟧ = $\lambda f_{\langle s,\, t\rangle}.\exists$e:f(e)
       d.   { ⟦ Asp_R v_{CAUSE} *the flowers bloomed* ⟧,
            ⟦ Asp_R v_{CAUSE} *the flowers blossomed* ⟧,
            ⟦ Asp_R v_{CAUSE} *the flowers wilted* ⟧, ... }
            = { $\exists$e,e′:bloomed(*the.flowers*′, e′) $\wedge$ CAUSE(e, e′),
               $\exists$e,e′:blossomed(*the.flowers*′, e′) $\wedge$ CAUSE(e, e′),
               $\exists$e,e′:wilted(*the.flowers*′, e′) $\wedge$ CAUSE(e, e′), ... }

The semantic interpretations, then, are that something caused the flowers to enter a bloomed, blossomed, or wilted state.

6.   Conclusion

Following an explanation of how to distinguish resultative passives from eventive and stative passives, I explored the idea that, contra recent claims by Meltzer-Asscher (2012), McIntyre (2013), Bruening (2014), and Alexiadou et al. (2015), resultative passives must be causative and non-agentive, and showed how the possible readings of resultative passives support this analysis. I used this approach to explain the seemingly contradictory results of diagnostics for agentivity, given a notion of what a possible causer can be that is restricted by encyclopedic knowledge. I also showed a set of diagnostics for causativity that can be used for unaccusative verbs. These diagnostics show that causative unaccusative verbs can productively form resultative passives while non-causative unaccusative verbs cannot. Taken together, these allow for a unified analysis of the resultative passive of transitive and unaccusative verbs.

   More broadly, this analysis prompts a reconceptualization of what the passive is. For example, Legate (2017) analyzes the passive as 'demotion of the thematic subject...through existential closure applying on the Voice head that introduces the thematic subject itself.' She further proposes that her analysis correctly predicts a *lack* of unaccusative passives. However, as shown

here, unaccusatives can productively form resultative passives, provided that existential closure can apply to a causing event. We can therefore consider the passive as a general process of existential closure applying to *any* kind of external argument. Differences between eventive and resultative passives could reduce to what kind of external argument undergoes existential closure: an agent or a causing event.

Michael Wilson
University of Massachusetts Amherst
mawilson@linguist.umass.edu

References

Alexiadou, A. (2014). The problem with internally-caused change-of-state verbs. *Linguistics* 52:4, pp. 879–909.

Alexiadou, A. & F. Schäfer (2006). Instrument subjects are agents or causers. Baumer, D., D. Montero & M. Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, Cascadilla Proceedings Project, Somerville, MA, pp. 40–48.

Alexiadou, A., E. Anagnostopoulou & F. Schäfer (2006). The properties of anticausatives crosslinguistically. Frascarelli, M. (ed.), *Phases of Interpretation*, Mouton de Gruyter, Berlin, Germany, pp. 187–212.

Alexiadou, A., B. Gehrke & F. Schäfer (2014). The argument structure of adjectival participles revisited. *Lingua* 149, pp. 118–138.

Alexiadou, A., E. Anagnostopoulou & F. Schäfer (2015). *External Arguments in Transitivity Alternations: A Layering Approach*. Oxford University Press, Oxford, United Kingdom.

Anagnostopoulou, E. (2003). Participles and Voice. Alexiadou, A., M. Rathert & A. von Stechow (eds.), *Perfect Explorations*, Mouton de Gruyter, Boston, Massachusetts and Berlin, Germany, pp. 1–36.

Anagnostopoulou, E. (2013). Little v, Voice, and phases in participles. Unpublished Handout.

Baker, M., K. Johnson & I. Roberts (1989). Passive arguments raised. *Linguistic Inquiry* 20:2, pp. 219–251.

Bruening, B. (2014). Word formation is syntactic: Adjectival passives in English. *Natural Language and Linguistic Theory* pp. 363–422.

Deal, A. R. (2009). The origin and content of expletives: Evidence from "selection". *Syntax* 12:4, pp. 285–323.

Embick, D. (2004). On the structure of resultative participles in English. *Linguistic Inquiry* 35:3, pp. 355–392.

Folli, R. & H. Harley (2005). Flavours of v: Consuming results in Italian and English. Kempchinsky, P. & R. Slabakova (eds.), *Aspectual Enquiries*, Springer, Dordrecht, The Netherlands, pp. 95–120.

García-Pardo, A. (2017). Aspect and argument structure in adjectival passives. *Borealis: An International Journal of Hispanic Linguistics* 6:1, pp. 21–52.

García-Pardo, A. (To appear). The aspectual composition of adjectival passives. To appear in the Proceedings of 50th Meeting of the Chicago Linguistic Society.

Gehrke, B. (2015). Adjectival participles, event kind modification, and pseudo-incorporation. *Natural Language and Linguistic Theory* 33:3, pp. 897–938.

Gese, H., C. Maienborn & B. Stolterfoht (2011). Adjectival conversion of unaccusatives in German. *Journal of Germanic Linguistics* 23:2, pp. 101–140.

Koontz-Garboden, A. (2009). *States, Changes of State, and the Monotonicity Hypothesis*. Ph.D. thesis, Stanford University.

Kratzer, A. (2000). Building statives. Conathan, L. J., J. Good, D. Kavitskaya, A. B. Wulf & A. C. L. Lu (eds.), *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Aspect*, pp. 385–399.

Kratzer, A. (2002). The event argument and the semantics of verbs. Unpublished draft.

Kratzer, A. (2005). Building resultatives. Maienborn, C. & A. Wöllstein (eds.), *Event Arguments: Foundations and Applications*, Max Niemeyer Verlag, Tübingen, Germany, pp. 177–212.

Legate, J. (2017). On passives. Talk presented 3 March 2017 at Harvard University.

Levin, B. (2009). Further exploration of the landscape of causation: Comments on Alexiadou and Anagnostopoulou. Halpert, C., J. Hartman & D. Hill (eds.), *Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT*, MIT Working Papers in Linguistics, Cambridge, Massachusetts, vol. 57.

Levin, B. & M. Rappaport Hovav (1995). *Unaccusativity: At the Syntax-Lexical Semantics Interface*. MIT Press, Cambridge, Massachusetts.

Maienborn, C. (2009). Building event-based *ad hoc* properties: On the interpretation of adjectival passives. Riester, A. & T. Solstad (eds.), *Proceedings of Sinn und Bedeutung 13*, pp. 31–45.

McIntyre, A. (2013). Adjectival passives and adjectival participles in English. Alexiadou, A. & F. Schäfer (eds.), *Non-Canonical Passives*, John Benjamins, pp. 21–41.

McKoon, G. & T. Macfarland (2000). Externally and internally caused change of state verbs. *Language* 76:4, pp. 833–858.

Meltzer-Asscher, A. (2012). Verbal passives in English and Hebrew: A comparative study. Everaert, M., M. Marelj & T. Siloni (eds.), *The Theta System: Argument Structure at the Interface*, Oxford University Press, Oxford, United Kingdom.

Partee, B. (1983/1997). Uniformity vs. versatility: The genitive, a case study. van Benthem, J. & A. ter Meulen (eds.), *The Handbook of Logic and Language*, Elsevier, pp. 464–470.

Pesetsky, D. (1995). *Zero Syntax: Experiencers and Cascades*. Current Studies in Linguistics, MIT Press, Cambridge, MA.

Pylkkänen, L. (2002). *Introducing Arguments*. Ph.D. thesis, MIT.

Pylkkänen, L. (2008). *Introducing Arguments*. MIT Press, Cambridge, Massachusetts.

Ramchand, G. (2008). *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge University Press, Cambridge.

Roeper, T. (1987). Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18:2, pp. 267–310.

Schäfer, F. (2007). By itself. Ms., Universität Stuttgart.

Schäfer, F. (2012a). The passive of reflexive verbs and its implication for theories of binding and case. *Journal of Comparative Germanic Linguistics* 15:3, pp. 213–268.

Schäfer, F. (2012b). Two types of external argument licensing – the case of causers. *Studia Linguistica* 66:2, pp. 128–180.

Sleeman, P. (2014). From participle to adjective in Germanic and Romance. Sleeman, P., F. Van de Velde & H. Perridon (eds.), *Adjectives in Germanic and Romance*, John Benjamins Publishing Company, Philadelphia, Pennsylvania, pp. 171–198.

Wasow, T. (1977). Transformations and the lexicon. Culicover, P., T. Wasow & J. Bresnan (eds.), *Formal Syntax*, Academic Press, New York, New York, pp. 327–360.

Williams, E. (1974). *Rule Ordering in Syntax*. Ph.D. thesis, MIT.