Phonological effects on noun-adjective ordering in Italian
A-Exam

Katherine Blake

25 May 2020

1 Introduction

A long-standing issue in theoretical linguistics is what the nature of the relationship is between the various sub-systems into which we have traditionally, and idyllically, organized language phenomena: syntax, semantics, phonology, and/or phonetics. Examining the syntax-phonology interface more closely, it was long thought that this relationship was a one-sided one: once the syntactic processes were carried out, this output became the input to a phonological spell-out. The syntactic structure was the string over which phonology then operated its own (separate) set of rules (Chomsky, 2010). An ever-growing body of literature (most notably since Inkelas and Zec, 1990), however, instead suggests a bi-directional relationship between these two systems. Not only is there compelling evidence for structural influence on phonology, e.g., through prosodic structure, but for phonological influence on structure as well, e.g., Heavy NP shift. While there is a large body of literature showing evidence of syntactic influences on phonology, significantly less work has been done showing evidence of phonological influences on syntax. This paper contributes findings of phonological effects on syntactic processes as evidence for this close relationship.

In order to examine the nature of the syntax-phonology interface, namely in the direction of phonology → syntax, the potential effects of phonological markedness avoidance on the

syntactic ordering of {noun, adjective} pairs in Italian were investigated. We know that many phonological phenomena are avoided by many unrelated languages, such as adjacency of two vowels across words or morphemes (vowel hiatus) and adjacency of two prominent syllables (stress clash) or otherwise non-alternating prominences, among others. Evidence for these phenomena being "marked" or dispreferred comes from repair strategies that we observe in surface forms that mitigate the surfacing of dispreferred underlying forms created by morpheme or word concatenation. An Italian example of a phonological repair of underlying vowel hiatus created by word concatenation is shown in (1).

(1) 'and here' UR: /e ekko/ → SR: [e**d** ekko] *ed ecco*

In (1), a [d] is phonologically inserted to repair the underlying /e e/ sequence created by the syntactic sequence [e ekko]. This paper examines phonologically marked phenomena created by word concatenation, to see whether they are repaired or avoided syntactically.

To do this, patterns of the internal ordering of {noun, adjective} pairs from an Italian corpus were analyzed as a function of the occurrence of various phonologically-marked structures: stress clash, vowel hiatus, and phonological weight (i.e., word length). Italian, and all other major Romance languages to varying degrees, has a set of adjectives which can appear in either position relative to the noun: [noun adjective] post-nominal or [adjective noun] pre-nominal. Post-nominal [noun adjective] order is considered the syntactic default. The adjectives which have some free variation are referred to here as flexible adjectives, and they often belong to certain semantic classes like beauty, age, goodness, and size. For these flexible adjectives, though variable, the default order is considered pre-nominal. See an example in (2) below.

(2) *bella città* ∼ *città bella* 'beautiful city'

The central hypothesis of this work is that the ordering of {noun, flexible adjective} pairs is manipulated by speakers to avoid phonologically-marked phenomena. Here the effects of stress clash, vowel hiatus, and light-final pairs are investigated. Reordering can sometimes

cause a slight change in the semantic meaning of the pair, so this potential factor is discussed as well, but is not a focus of the paper.

2 Background

2.1 Phonological Effects on Syntax

In the phonological literature, the proposed interaction between phonology and syntax is nothing new. Inkelas and Zec (1995) review several cross-linguistic phenomena that serve as evidence for the bi-directional connection between phonology and syntax. It is no longer the case that a Y-model in which syntax simply feeds phonology can be assumed. There has to be a mutually-influential relationship. This notion is argued by Inkelas and Zec (1995) using evidence from both phonology and syntax of the other domain's influence. They review phonological phenomena shown to be sensitive to syntactic structure including *Raddoppiamento Sintattico* in Italian and stress retraction in English (see section 2.4). These are two processes which occur at word boundaries, but only when there exists a certain syntactic relationship between the two words. This is to show the influence of syntax on phonology.

More crucial to this work is the evidence presented for phonological influence on syntax. Inkelas and Zec (1995) present data from Serbo-Croatian and English as evidence (as described in Inkelas and Zec, 1990). In Serbo-Croatian, topicalization, a syntactic process, can occur only with constituents that are at least a branching phonological phrase (as in (3a)). Phrases which consist of a topicalized constituent that is only one phonological word (as in (3b)) are judged to be ungrammatical.

In English, data in support for phonological influence on syntax come from a well-known phenomenon called Heavy NP Shift. In constructed which have a heavy NP, this "shifted" phrase is minimally two phonological phrases (as in (4a)). If the shifted NP is lighter than this minimum, the sentence is judged to be ungrammatical (as in (4b)).

(3) a. [[Petar]$_\omega$ [Petrović]$_\omega$]$_{NP}$ voleo-je   Mariju
     Peter Petrovic   loved-AUX   Mary
     'Peter Petrovic loved Mary'
   b. *[[Petar]$_\omega$]$_{NP}$ voleo-je   Mariju
     Peter   loved-AUX   Mary
     'Peter loved Mary'

(4) a. Mark showed to John   [[some letters]$_\phi$ [from Paris]$_\phi$]$_{NP}$
   b. *Mark showed to John   [[some letters]$_\phi$]$_{NP}$

There are two key works that inform the present study: one is an instance of phonological effects on morphology and the other is on syntax. Martin (2011) examines morpheme-internal phonotactic constraints in Navajo (required sibilant harmony) and English (ban on geminate consonants). In both languages, these constraints are active within morphemes, but are violated across prosodic word boundaries like compounds. For example, a geminate is not allowed within a root in English, but there do exist compounds like *roommate*. While these are legal, Martin shows them to be statistically underrepresented. He proposes this is due to the absence of monomorphemes that have geminates, and since a learner constructs a grammar based on their linguistic input, they will underpredict the number of compounds with geminates because they generalize the rule for monomorphemes to heteromorphemic words.

Additionally, Breiss and Hayes (2019) provide evidence for phonological effects on syntax in their work on English. They examine phonological markedness effects on sentence formation by looking at the avoidance of several phonological markedness constraints (including stress clash, as is examined in this paper) in sentence bigrams. Using a MaxEnt model to diagnose these avoidances, it was found that syntactic choice (i.e., word order) and lexical choice (i.e., synonym selection) were significant strategies for phonological markedness avoidance.

A study similar to the present one is presented in Shih and Zuraw (2017). The authors use a large corpus to examine noun-adjective ordering in Tagalog. Nouns and adjectives can either appear as noun-linker-adjective, or adjective-linker-noun (the default). The linker can

appear as *-ng* or *-na*, and the variation is phonologically conditioned. Various phenomena were tested using corpus data, and the authors found that several phonological effects were active in determining word order of adjective-noun. Given that *-ng* ends in a nasal and nasals are allowed in onset position in Tagalog, it was found that obeying the OCP for nasals (i.e., avoiding a nasal-nasal sequence) was favored over default ordering of adjective-noun. The avoidance of a nasal before a voiceless consonant was also found to be a factor in word order, as was the avoidance of vowel-vowel sequences (i.e., *-na* before a word with no onset). This additional evidence from Tagalog of phonological effects (both segmental and syllabic) on adjective-noun ordering in the language strengthens the case for the influence of phonology on syntax.

These phonological effects on syntax were tested in languages like Navajo, English, and Tagalog. In Italian and other Romance languages, much work has been done showing evidence of syntactic effects on phonology (stress shift in Italian (Nespor and Vogel, 1986); focus in Italian (Frascarelli, 2000); intonational/accentual phrasing in French (Jun and Fougeron, 2000); intolerance of vowel hiatus in Spanish (Varis, 2012)). Not much work has been done, however, to show the reverse: phonological effects on syntax.

2.2   Syntactic Framework

The syntactic framework used in this paper is greatly informed by Cinque (2010). In his book, he outlines two different types of adjectives, direct modification (DM) and indirect modification (IM). DM adjectives come from a functional projection and are non predicative. On the other hand, IM adjectives come from reduced relative clauses and are predicative. Cinque proposes that DM adjectives can be pre or postnominal while IM adjectives are only postnominal in Italian. The order is therefore as follows: AP_{DM} NP AP_{DM} AP_{IM}. The derivation for this ordering is shown in Figure 1.

In Figure 1, the NP can optionally raise over DM adjectives, which have flexible ordering. If (1) does not occur, DM adjectives are prenominal; if (1) does occur, DM adjectives are
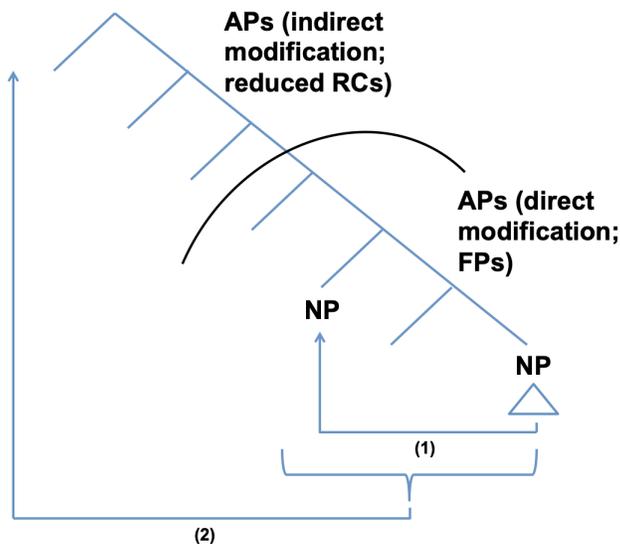
Figure 1: Derivation of adjective ordering in Italian based on Cinque (2010)

postnominal. Then, in (2), the NP and the AP_{DM} together raise above the AP_{IM}. (2) is not optional, as AP_{IM} is always post-nominal.

While DM adjectives are flexible, Cinque claims they have a strict semantic interpretation when in prenominal position and an ambiguous interpretation when in postnominal position (between their prenominal/DM meaning, and the meaning associated with postnominal/IM adjectives). The semantic meanings associated with each type of adjective in Italian are in Figure 2, a table from Cinque (2010), p. 17. These semantic distinctions are touched upon in this paper, but are not the focus of this work.

2.3   Italian Adjectives

This study examines only qualifying or lexical adjectives (descriptors), and does not include an investigation of determinative adjectives (such as possessives, demonstratives, indefinites, and interrogatives). This is because determinative adjectives are generally more restrictive (though there is some flexibility with possessives, which I will not get into here.) Lexical adjectives show morphological agreement in gender and number with the nouns they modify. Gender and number agreement morphology has some variation, but all adjectives end in one

Prenominal adjectives	N	Postnominal adjectives
individual-level reading		individual-level or stage-level reading
nonrestrictive reading		restrictive or nonrestrictive reading
modal reading		modal or implicit relative clause reading
nonintersective reading		intersective or nonintersective reading
absolute reading		relative or absolute reading
absolute reading of superlatives		comparative or absolute reading of superlatives
specificity-inducing reading		specificity- or non-specificity-inducing reading
evaluative reading of 'unknown'		evaluative or epistemic reading of 'unknown'
NP-dependent reading of 'different'		NP-dependent or discourse anaphoric reading of 'different'

Figure 2: Semantic interpretations associated with pre-nominal (DM) and post-nominal (DM or IM) adjectives in Italian from Cinque (2010) p. 17

of the following vowels: /o/ typically masculine singular, /i/ masculine plural, /a/ feminine singular, or /e/ feminine plural[1].

Not all adjectives in Italian can occur in both prenominal and postnominal positions. Only adjectives which can have both a DM and an IM semantic reading can occur in both positions (Figure 2), and it is only in postnominal position that both these readings are available. An example of the flexible adjective *buono* 'good', which can have both intersective and nonintersective readings, is shown below in (5); this is in constrast with an adjective like *ex* 'former' with only one reading, which cannot occur in both positions, shown in (6) (data from Cinque, 2010).

To my knowledge, there has been no previous work regarding phonological effects on adjective ordering in Italian; though, some work in this vein has been done on other languages (review Section 2.1).

---

[1]Except for some adjectives which have invariable morphology.

(5) a. Un **buon** attaccante non farebbe mai una cosa del genere.
     a   good   forward   not   would-do   never   a   thing   of-the   kind
     'A forward good at playing forward would never do such a thing' (*nonintersective*)
   # 'A good-hearted forward would never do such a thing' (*intersective*)
   b. Un attaccante **buono** non farebbe mai una cosa del genere.
     a   foward   good   not   would-do   never   a   thing   of-the   kind
     'A forward good at playing forward would never do such a thing' (*nonintersective*)
     'A good-hearted forward would never do such a thing' (*intersective*)

(6) a. l' **ex** primo ministro è stato arrestato
     the former prime minister has been arrested
   b. *il primo ministro **ex** è stato arrestato
     the prime minister former has been arrested
     'The former prime minister has been arrested.' (*strictly nonintersective*)

2.4 Italian Phonology

2.4.1 Stress Clash

Stress clash is a phonological phenomenon that is defined as two prominent syllables occurring adjacent to each other. Two examples are given in (7), one in English and one in Italian.

(7) a. thirteen men   *thirtéen mén*
   b. città vecchia   *cittá vécchia*   'old city'

The English example is included because it is an oft-cited example showing *stress retraction*. Stress retraction often occurs to avoid stress clash, whereby the stressed syllable which is not the primary stress of the phrase (in both English and Italian, this is the left stressed syllable occurring word-finally in word 1) retracts or is produced on the previous syllable instead. This yields *thírteen mén* on the surface as opposed to a faithful realization of the two underlying forms concatenated *thirtéen mén*. Nespor and Vogel (1979) investigate this same phenomenon in Italian. Their experimental work on speakers of standard northern Italian revealed that those speakers behaved similarly to the English example above, retracting stress from word 1 back a syllable when word 2 had stress on the first syllable (i.e., there was a stress clash at the boundary between word 1 and word 2). They found this was

more likely to occur when word 1 and word 2 belonged to the same syntactic phrase, and more specifically when word 2 was the head of that phrase, at the right edge. This previous finding supports the prediction in the present study that underlying [adjective noun] pairs are more likely to be sensitive to stress clash than [noun adjective] pairs.

Nespor and Vogel (1979) discuss another strategy Italian speakers employ to avoid stress clash, *Raddoppiamento Sintattico* (RS). Rather than retract stress to the previous syllable speakers who use RS instead lengthen the initial consonant of word 2. An example is shown in (8).

(8) città vecchia   *cittá vécchia* → *cittá <u>vv</u>écchia*   'old city'

The geminate onset – as opposed to the underlying singleton onset – is thought to create enough distance between the two stresses to remedy the disfavored stress clash occurrence. Nespor and Vogel (1979) found that speakers employ either stress retraction or RS as their avoidance strategy, but never both. Stress retraction varieties do also exhibit lengthening, but of the final vowel in word 1 in an attempt to also pad the distance between the two prominent syllables if stress clash must occur. This is often referred to as the rhythm rule.

Both of these strategies, however, are considered optional by Nespor and Vogel. Where they are not employed and stress clash does surface, these instances have been described as "real-life situations" (Liberman and Prince (1977)). Nespor and Vogel (1979) push back against this, instead accounting for allowances of stress clash using syntactic groupings. If word 1 and word 2 are not sisters of a syntactic phrase, or word 1 is the head, then they argue stress clash is not as likely to be avoided. Corpus results in the present study, however, indicate that even in {adjective noun} pairs, stress clash still surfaces. I will argue that in these instances, it is possible that the "real-life situations" Liberman and Prince (1977) refer to may still hold, and can be defined more specifically as pragmatic or semantic motivations for allowing stress clash in order to preserve or emphasize a particular meaning associated with pre-nominal or post-nominal ordering.

Finally, it is clear that Nespor and Vogel (1979) operate under the assumption that

the syntactic output serves as the input material for phonology: "When clash arises in a given string, a phonological rule operates to eliminate it" (p. 476). This paper argues for a mutual influence between syntax and phonology, in which case phonological output is known by syntax, and therefore syntactic operations such as word order can also be employed as strategies for the avoidance of stress clash, and other disfavored phonological phenomena.

2.4.2 Vowel Hiatus

The second marked phonological phenomena investigated in this work is vowel hiatus. Hiatus occurs when two vowels are adjacent to each other, and they belong to separate syllables. A word-internal examples from Italian are shown in (9) below, where in (9a) the high vowel becomes a glide, but nothing is done to repair hiatus in (9b) (data from Kramer, 2009, p.52).

(9) a.   buono   [ˈbwɔ.no]   'good'
     b.   paura   [pa.ˈuː.ɾa]   'fear'

In general, typologically-unrelated languages tend to disprefer instances of vowel hiatus and it is common for a language to have some sort of repair strategy that either deletes one of the offending vowels, or inserts an epenthetic consonant to break up the sequence. An example from Italian shown in section 1 is repeated below.

(10)   'and here'   UR: /e ekko/ →   SR: [e**d** ekko]   *ed ecco*

The [d] surfaces to prevent an [e.e] sequence, which would be an instance of vowel hiatus. This kind of repair in Italian, however, is not common. (10) is one example, and there are only a few other fixed expressions in which a similar epenthetic process is found.

On the other hand, like other Romance languages, there is vowel elision between vowel-final articles and vowel-initial nouns. An example from Italian is shown in (11).

(11)   'the university (fem)'   UR: /la u.ni.ver.si.ta/ →   SR: [lu.ni.ver.si.ta]   *l'università*

This relatively mixed tolerance-level of hiatus is in stark contrast to French, another Romance language, which has quite a complex and pervasive vowel hiatus repair phenomenon,

liaison. In Italian, epenthetic consonants appear in some fixed expressions to avoid it, and it is avoided via vowel deletion between clitics and nouns. The clitic-noun relationship could be a case of exceptionality due to the particular dependence of clitics on nouns, or weak faithfulness; but, vowel hiatus is often permitted as exemplified by (9) above, given the right circumstances[2]. Because of this, results for syntactic avoidance of hiatus in this study are expected to be mixed as well. This is in contrast to the predictions for stress clash. Since clash has been shown to be active in Italian phonology, it is predicted to be actively avoided by means of word-order manipulation.

2.4.3   Phonological Weight

As pointed out in Inkelas and Zec (1990), a clear effect of phonology on syntax can be found when we look at the organization of a sentence given constituents of different lengths[3]. Similar to Heavy NP shift noted for datives in English (Inkelas and Zec, 1995), a similar phenomenon is present in Italian for subjects.

Italian is typically described as an SVO language, but other sentence structures are possible and exploited for emphasis or artistic motivations. For example, the sentence in (12a) is in the unmarked order, SVO; however, the OVS order shown in (12b) is also grammatical. Data in (12) and (13) are from Cardinaletti (2010).

(12) a. *Il   partito   di   maggioranza   fece   poi   la   stessa   proposta.*
          the   party   of   majority   made   then   the   same   proposal
      b. *La   stessa   proposta fece   poi   il   partito   di   maggioranza.*
          the   same   proposal made   then   the   party   of   majority
          'The majority party then made the same proposal (not a similar one)'

This OVS structure is ungrammatical, however, when the subject NP is light (only one phonological word). This is shown in (13).

---

[2]Such as lexical stress assignment to the second vowel in a V.V sequence, and the particular quality of the two vowels. See Kramer (2009) for more details.

[3]While it is common in the phonological literature to use *weight* in reference to mora, i.e., to refer to the segmental level, in this paper I use weight to refer to word length.

(13)   *La stessa proposta fece poi Gianni/lui.

Given this process of Heavy NP shift in Italian, here involving heavy subjects, it is predicted more generally that the structure within a constituent such as the {noun, adjective} pair, where order is flexible, is sensitive to the weight of an NP, just as it is in the case of SVO → OVS word order. Though in the noun-adjective case in this study, it operates over a smaller scale (number of syllables instead of number of prosodic words), the constituents over which this rule is predicted to operate are also at a smaller scale, within an NP.

3 Methodology

3.1 Hypotheses and Predictions

The principal hypothesis of this work is that the interface between syntax and phonology is bi-directionally influential. It is the goal of this work to show evidence for phonological influence on syntax, as much work has already shown the effects of syntax on phonology. It is, however, important to note a crucial stipulation. While it is hypothesized that phonology influences syntax, it is not the case that *all* phonological processes can have this effect. As pointed out in Shih and Zuraw (2017), only the phonological processes that are *active* in a language are predicted to have some influence over the syntax. It not logical to hypothesize that a phenomenon not otherwise present in a language's phonology, even if it is considered marked by phonologists given cross-linguistic evidence, will see ripple effects in that language's syntax.

Thus, the hypothesis and following predictions are outlined below.

Hypothesis: Phonologically-marked phenomena in a language that are avoided phonologically, will also be avoided syntactically, where possible.

(a) **Prediction 1 (Stress Clash)**: Given that stress clash is actively avoided in Italian phonology, where word order is flexible, it will also be avoided syntactically

via word-order manipulation in {noun, adjective} pairs.

(b) **Prediction 2 (Vowel Hiatus)**: Given that vowel hiatus can be said to *not* be actively avoided in Italian phonology, because its tolerance level is so mixed, even where word order is flexible, it will not be avoided syntactically via word-order manipulation in {noun, adjective} pairs.

(c) **Prediction 3 (Phonological Weight)**: Given that syntactic structure is sensitive to the phonological weight of NPs elsewhere in Italian (e.g., in OSV ordering), it will also be sensitive to phonological weight in flexible {noun, adjective} pairs.

3.2 Corpus

Data to test the above hypothesis come from the Universal Dependencies corpus of Italian located at github.com/UniversalDependencies/UD_Italian-ISDT. The text corpus contains 14,167 sentences, 278,429 tokens, and 298,344 words. It is made up of several corpora, and contains a mixture of written and oral text of various registers. The breakdown of the corpus is shown in Figure 3.

Each word in the corpus has word, lemma, part of speech, and morphological (including gender, number, definiteness, type of pronoun, clitic status, mood, person, tense, verb form) information. Sentences are demarcated and their original source is coded in their ID, thus making it possible to trace their status as written versus oral. This distinction could certainly play a role in the effects of phonology on syntax, but is left to future work.

Any noun followed by an adjective and any adjective followed by a noun constituted a {noun, adjective} pair. Adjectives were categorized as *flexible* if they appeared in both pre and postnominal positions somewhere in the corpus.

Original format	Source	Genre	Size in tokens	Size in sentences
TUT-CONLL	Evalita 2011 Dependency parsing	Legal texts, news articles, Wikipedia articles	101,309	3,842
ISST-TANL	Evalita 2011 Domain adaptation task	Newspaper articles	80,967	4,135
ISST-TANL	SPLeT 2012	Legal texts: European directives	6,166	260
MIDT	Several QA competitions	Questions	20,680	2,228
MIDT	Evalita 2014 Dependency parsing:test data set (partial)	News articles	7,618	304
TUT-CONLL	Parallel TUT (Italian part)	Various genres	55,942	2,131
UD	Due Parole	Simplified Italian news	24,977	1,421
UD2	New data	Various sentences	2,504	150

Figure 3: Composition of the UD Italian corpus

3.3 Lexical Database

There is no phonological information in the UD corpus, so the PhonItalia lexical database was utilized for stress and number of syllable information. PhonItalia is an open-source database with phonological information for 120,000 word forms of Italian (Goslin et al. 2013). Once the {noun, adjective} pairs were collected from the UD corpus, their phonological information was looked up in PhonItalia. If the word form did not appear in the lexical database, the pair was thrown out. With stress and syllable count information from PhonItalia, clash and relative phonological weight were calculated between the two members of each pair; with phone information, hiatus was calculated (Table 1).

14

Categorization	Definition
Unknown clash	Stress information from PhonItalia was not available for one member of the pair
Unavoidable clash	Both members of the pair are monosyllablic
Clash	There is stress clash between the pair, in the order in which they appear in the corpus, i.e., member 1 has ultimate stress and member 2 has initial stress
Avoided clash	If the members of the pair were in the opposite order from which they appear in the corpus, there would be a stress clash between them, i.e., member 2 has ultimate stress and member 1 has initial stress
Impossible clash	There is no possible stress clash between the two members of the pair (even with reordering), e.g., member 1 or member 2 has penultimate stress
Unknown hiatus	Phone information from PhonItalia was not available for one member of the pair
Unavoidable hiatus	Both members of the pair have vowels at both the left and right edges
Hiatus	There is vowel hiatus between the pair, in the order in which they appear in the corpus, i.e., member 1 has a vowel at the right edge and member 2 has a vowel at the left edge
Avoided hiatus	If the members of the pair were in the opposite order from which they appear in the corpus, there would be vowel hiatus between them, i.e., member 2 has a vowel at the right edge and member 1 has a vowel at the left edge
Impossible hiatus	There is no possible vowel hiatus between the two members of the pair (even with reordering), e.g., member 1 or member 2 both have a consonant at the left edge
Light-final	Instances where member 2 has fewer syllables than member 1, as they appear in the corpus
Heavy-final	Instances where member 2 has *more* syllables than member 1, as they appear in the corpus
Equal-weight	Instances where members of the pair have the same number of syllables

Table 1: Phonological information for clash, hiatus, and weight collected based on PhonItalia

3.4 Statistical Tests

3.4.1 Monte Carlo Procedure

Rates of the various types of clash, hiatus, and weight are analyzed in this work. In order to determine whether the distributions of these different phenomena differ significantly from chance, the Monte Carlo Procedure was used (Martin (2011)). This procedure approximates the expected distribution of light-final vs. heavy-final pairs, for example, if they were more randomly combined. This is done by taking all nouns and all adjectives and randomly combining them as pairs in random order (1000 samples, with replacement; 100 runs). The rate of prenominal pairs in the reshuffled data was set to 50% to maintain randomness. The reshuffled data included only flexible adjective pairs in which the phonological phenomenon

being tested was relevant. Continuing with the relative heaviness example, the Monte Carlo procedure involved randomly reshuffling nouns and adjectives from all flexible pairs which were either light-final or heavy-final in the corpus, ignoring equally-weighted pairs. The rate of light- vs. heavy-final is then calculated from the Monte Carlo set and compared with that found in the true dataset (also excluding all equal-weight pairs). This reshuffling was repeated several times to ensure that the rate produced by the Monte Carlo Procedure was fairly consistent across calculations, and thus a reliable measure. With the Monte Carlo distribution, it is then possible to calculate how likely the trends found in the actual data are. This was determined using a one-sample T-test.

From the predictions stated above, I expected to find rates of avoided clash and heavy-final pairs to be higher in the real data than in the Monte Carlo distributions, showing that what speakers are doing (actual data) is avoiding these marked phenomena more than what would happen if nouns and adjectives combined more freely (Monte Carlo data). The prediction for vowel hiatus is that it would be unclear, so the expected result from the Monte Carlo Procedure is that the real data and the Monte Carlo distribution are not significantly different. All three of predictions were confirmed, and are reported in the Results section.

3.4.2   Regression models

Mixed-effects logistic regression models run using R reported in the next section were run using the three phonological effects as fixed effects, and words categorized as adjectives as the random effect, in order to predict word order (Core Team et al., 2013). Adjectives were chosen as a random effect in case there were certain adjectives driving the effects of clash, hiatus and/or weight (i.e., a set of high-frequency adjectives with final stress driving the clash effect). Nouns were not chosen as a random effect because there was a larger unique set of them, and linguistically it is the identity of the adjective that makes the order of a {noun, adjective} pair flexible due to its semantics, rather than the identity of the noun.

Logistic regression models are also reported using the same fixed effects of the three

phonological factors, but to predict adjective flexibility (Core Team et al., 2013). In these models, the random effect of adjective is not included because of the desire for the model to rely on the phonological effects to predict the type of adjective, rather than certain adjectives themselves. The lexical items which are adjectives was deemed too overlapping with what was being predicted.

How the fixed effects and the categorical variables being predicted were coded is shown in Table 2. In general, fixed effects were coded as 1 if the marked phonological phenomenon was avoided or impossible, and 0 if it actually occurred in the corpus. Because of this, a positive coefficient value in the results for a phonological factor would indicate that prenominal order or adjective flexibility is predicted by avoidance of that phonological phenomenon. This is what is predicted to be found for clash and weight from the hypothesis in 3.1. The models were run in order to see if default order was overridden by the desire to avoid a phonologically-marked sequence in that order; and, if occurrences of avoidance of these phonologically-marked sequences were associated with adjective flexibility. This method is a two-fold way to confirm my hypothesis that phonology has an effect on syntax in these data. These predictions are generally confirmed, and are reported in the Results section.

Fixed effect	Value of 1	Value of 0
Clash Factor	if avoided or impossible clash	if true clash
Hiatus Factor	if avoided or impossible hiatus	if true hiatus
Weight Factor	if heavy-final or equal-weight	if light-final
Prediction Variable	Value of 1	Value of 0
Order	if prenominal (non-default)	if postnominal (default)
Adjective flexibility	if flexible	if strictly prenominal or strictly postnominal

Table 2: How variables for regression models are coded

4   Results

This section is formatted as follows:

Section 4.1 presents some descriptive statistics of the data: overall, and for each phonologically-marked phenomenon.

Section 4.2 presents the results of a mixed-effects logistic regression, predicting {noun, adjective} order from the three phonological factors: clash, hiatus, and weight. For this model, it is predicted that avoided clash and heavy-final pairs correlate with prenominal order; meaning, clash and light-final occurrences are predicted to push the order from the default (postnominal) to the more marked (prenominal). Hiatus was predicted to not necessarily have an effect on order. In the same section are the results of a logistic regression predicting adjective flexibility from the three phonological factors. For this model, it is predicted that avoided clash and heavy-final pairs predict a flexible adjective; meaning, avoidance of phonological markedness is correlated with flexible word order. Hiatus was predicted to not necessarily have any predictive power on the flexibility of adjectives.

Section 4.3 presents results of the Monte Carlo Procedure for clash, hiatus, and weight. Rates of clash and light-final pairs are predicted to be higher in the Monte Carlo distributions for these phenomena; weight is predicted to not be significantly different from its Monte Carlo distribution.

Section 4.4 reports results on clash and hiatus occurrences with the {noun, adjective} pair "neighbors," meaning the directly left-adjacent and right-adjacent words, in order to probe effects of constituency on the tolerance level of these phonologically-marked phenomena. It is expected that clash, but not necessarily hiatus, occurs at higher rates with the neighbors, with which they may not form a syntactic constituent, in comparison to within the pairs.

A summary of the results can be found in 4.5.

4.1 Descriptive Statistics

4.1.1 Overview

Among the 14,498 {noun, adjective} pairs collected from the UD corpus, 4,689 adjectives appear only in prenominal position (12%); 9,809 only in postnominal position (60%); and 5,536 adjectives appearing in both positions (28%). 593 of the strictly prenominal adjectives are unique, 2,911 of the strictly postnominals are unique, and 1,364 of the flexibles are unique. In general, across all adjective types, it was found that 68% of all {noun, adjective} pairs in the corpus were postnominal, confirming the assertion in the literature that the default order is in fact [noun adjective].

4.1.2 Clash

Across all adjective types, occurrences of stress clash within pairs remained low. This is not surprising: though Italian stress is not always easily predicted, the penultimate syllable is generally preferred (Borrelli, 2013). Since stress is typically not placed at a word edge, clash is usually impossible. In regards to the hypothesis investigated in this work, the more interesting result is the trend of avoided clash versus allowed clash. Among the flexible pairs, results shown in Table 3, there were more instances of avoided clash pairs than true clash pairs.

Flexible adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	180	3%
Unavoidable	0	0%
Clash	80	1%
Avoided clash	108	2%
Impossible clash	5,168	93%
Total	5,536	

Table 3: Clash distribution results for flexible adjectives

This distribution among the flexible pairs where clash was possible (true clash vs. avoided

clash) is the comparison for the Monte Carlo Procedure presented in 4.3. The clash results for strictly prenominal and postnominal adjectives are shown in Tables 4 and 5.

Prenominal adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	172	11%
Unavoidable	4	<1%
Clash	53	3%
Avoided clash	30	2%
Impossible clash	1,274	83%
Total	**1,533**	

Table 4: Clash distribution results for strictly prenominal adjectives

Postnominal adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	613	8%
Unavoidable	0	0%
Clash	16	<1%
Avoided clash	25	<1%
Impossible clash	6,775	91%
Total	**7,429**	

Table 5: Clash distribution results for strictly postnominal adjectives

Among the pairs where order is not flexible, the trend of true vs. avoided clash was the opposite for prenominal (i.e., more instances of true clash than "avoided") but the same for postnominal (i.e., more instances of "avoided" clash than prenominal). Though if flexibility is compared (i.e., prenominal and postnominal results are collapsed), flexible adjectives are more likely to occur in a position where clash is avoided compared to the rate of clash found in fixed adjectives. This is a preliminary confirmation of my prediction, that flexibility increases the rate of avoided clash.

4.1.3 Hiatus

Across all adjective types, occurrences of vowel hiatus within each pair remained low. Instances of avoided hiatus also numbered very few. This was not unexpected, as words in

Italian are typically have onsets and often end in vowels. Among the flexible pairs, results shown in Table 6, there was virtually an equal amount of true hiatus pairs and avoided hiatus pairs.

Flexible adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	434	8%
Unavoidable	182	3%
Hiatus	794	14%
Avoided hiatus	790	14%
Impossible hiatus	3336	60%
Total	5,536	

Table 6: Vowel hiatus distribution results for flexible adjective pairs

This distribution among the flexible pairs where hiatus was possible (true hiatus vs. avoided hiatus) is the comparison for the Monte Carlo Procedure presented in 4.3. The hiatus distributional results for strictly prenominal adjectives and postnominal adjectives are shown in Tables 7 and 8.

Prenominal adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	214	14%
Unavoidable	83	5%
Hiatus	149	10%
Avoided hiatus	445	29%
Impossible hiatus	642	42%
Total	1,533	

Table 7: Vowel hiatus distribution results for strictly prenominal adjective pairs

Among the pairs where order is not flexible, the hiatus trends are more interesting. If a pair is strictly prenominal, it has a much greater chance of being an instance of "avoided" hiatus. The label "avoided" here is a bit misleading since the order of these pairs is not flexible, and therefore phonological processes cannot be avoided using word (re-)ordering; however, this could be the result of a diachronic process, rather than the synchronic trends

Postnominal adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	967	13%
Unavoidable	362	5%
Hiatus	1664	22%
Avoided hiatus	922	12%
Impossible hiatus	3514	47%
Total	7,429	

Table 8: Vowel hiatus distribution results for strictly postnominal adjective pairs

investigated in this work. The opposite trend is found for strictly postnominal pairs, where true hiatus is more likely than "avoided" hiatus, where hiatus is possible.

4.1.4 Weight

Previous literature has shown evidence for a preference of lighter phrases preceding heavier phrases (e.g., Heavy NP Shift, Inkelas and Zec, 1995); and of restrictions on OVS order in Italian due to heaviness of the subject NP (Cardinaletti, 2010). This is investigated here as well, within the NP. If Heavy NP shift is indicative of a more general preference, then where adjective placement is flexible, a greater likelihood for the heavier item (either the noun or the adjective) to come second in the pair is expected. Heavier is defined here as having a greater number of syllables than the other member of the {noun, adjective} pair. Results across adjective types are in Table 9, confirming this trend of heavy-final preference.

PHENOMENON	COUNT	PERCENTAGE
Light-final	3097	23.4%
Heavy-final	7386	50.9%
Equal weight	4015	27.7%
Total	14498	

Table 9: Weight results across all adjective types

Among pairs where the adjective is flexible, the trend is as predicted, with the plurality of pairs being heavy-final. This distribution is shown in Table 10. The distribution of light- vs. heavy-final flexible pairs is used as comparison to the Monte Carlo Procedure reported in

4.3. It is not the case, however, that a larger proportion of flexible adjectives were heavy-final when compared to strictly prenominal or postnominal adjectives, as was predicted.

Flexible adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Light-final	1,143	21%
Heavy-final	2,585	47%
Equal	1,808	33%
Total	5,536	

Table 10: Weight distribution results for flexible adjective pairs

Pairs which are strictly prenominal or postnominal showed the same trends as flexible: the plurality of both types is heavy-final. These distributions are shown in Tables 11 and 12.

Prenominal adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Light-final	355	23%
Heavy-final	754	49%
Equal	424	28%
Total	1,533	

Table 11: Weight distribution results for strictly prenominal adjective pairs

Postnominal adjectives		
{NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Light-final	1,599	22%
Heavy-final	4,047	54%
Equal	1,783	24%
Total	7,429	

Table 12: Weight distribution results for strictly postnominal adjective pairs

4.2 Regression Models

4.2.1 Predicting order

Using the `glmer` function from the `lme4` package in R, a mixed-effects logistic regression was run to see if the use of a non-default order of a {noun, adjective} pair could be predicted by any of the phonological effects investigated here (Core Team et al., 2013). Clash, hiatus, and weight were included as fixed effects and adjective identity as a random effect. The coefficients, standard deviations, and p-values for each of the fixed effects are reported in the table below (N = 5536).

Predicting: Order			
Fixed effect	Coefficient	SD	p-value
Clash Factor	0.59	0.26	p = 0.006*
Hiatus Factor	-0.65	0.10	p = 5.15e-10*
Weight Factor	-0.004	0.09	p = 0.96

Table 13: Mixed-effects logistic regression results

The model is interpreted as follows. Clash and hiatus were found to both be significant effects, meaning that clash and hiatus both help predict the surface order of a {noun, adjective} pair; however, weight was not significant, meaning it does not help predict order. The coefficient value for clash is positive, meaning that with the other factors held constant, as no correlation was found between the factors, occurrences of clash tend to push the order to be prenominal as to avoid clash[4]. The opposite was found to be true for hiatus. Since the coefficient for hiatus is negative, this is interpreted as the order remaining postnominal (default) in order to avoid hiatus occurring in the prenominal order, all other factors constant[5].

The results of this model confirm the prediction for clash: that adjectives are shifted prenominally (to the non-default position) in order to avoid clash. The results for hiatus

---

[4]Recall that order is coded as 1 if prenominal, 0 if postnominal. Clash is coded as 1 if avoided or impossible, and 0 if it occurred in the corpus.

[5]Recall that order is coded as 1 if prenominal, 0 if postnominal. Hiatus is coded as 1 if avoided or impossible, and 0 if it occurred in the corpus.

from this model were unexpected, and may be due to an historical effect. This, and why there was no effect found for weight, are discussed in more detail in the discussion section.

The following three subsections (4.2.2 - 4.2.4) report the outcome of the same mixed-effects logistic regression, but on different subsets of the flexible adjective data. This was done in order to investigate if avoidance of a phonological effect was truly driving the order, rather than a large set of pairs in which the phonological effect was not possible, since avoided and impossible were coded the same in the fixed effects variables. In general, for each smaller model, the relevant phonological factor had the same result as the full model. This means that removing impossible pairs for a particular phonological phenomenon did not have an effect on its predictive power: clash remained a significantly positive factor in the clash subset, hiatus was no longer significant in the hiatus subset, and weight remained insignificant, even in the weight subset. This is indicative of a more broad-scale effect of clash on the data, a general lack of effect of weight on the data, and mixed results for hiatus.

4.2.2 Predicting order: Clash subset

The same mixed-effects logistic regression model predicting order presented in 4.2.1 was run on a subset of the flexible adjective data that excluded all pairs were clash was impossible, in order to see if the coefficient for clash remained positive and significant when the large set of impossible clash data was not available to the model (Core Team et al., 2013). It is predicted that the coefficient for clash remains consistent, meaning that clash avoidance is correlated with a prenominal order, where flexible; and this is the result found by the model, reported in Table 14. Impossible pairs were a large part of the data, as shown in Table 3, so this brought the sample size down to N=188.

The results of the model are interpreted as follows. Clash and weight were both significant effects, meaning that they both help predict the surface order; however, hiatus is not a significant effect. The coefficient for clash is positive, showing the same effect as was found in the full model, suggesting adjectives are shifted prenominally in order to avoid clash.

25

Predicting: Order			
Fixed effect	Coefficient	SD	p-value
Clash Factor	4.91	0.31	p = 2.46e-09*
Hiatus Factor	-0.46	1.12	p = 0.68
Weight Factor	-1.2	0.44	p = 0.006*

Table 14: Mixed-effects logistic regression results: Clash subset

The results in this subset of the data, however, were different from the full model for the other two factors. Hiatus is no longer a significant effect; but weight was found to be significant, when it was not in the full model. Additionally, the coefficient for weight is negative, indicating that the order of {noun, adjective} pairs tends to remain postnominal in order to avoid a light-final pair caused by a prenominal order.

Upon closer inspection of this dataset, it appears that hiatus was no longer a significant effect because in this small subset of data, for most pairs hiatus was impossible. The unexpected negative weight effect is believed to be due to most of the pairs in this data subset being equal-weight or light-final (111 out of 188 pairs, 74 of those pairs being equal). Why most of these pairs happen to be equal weight, or the possibility that a light-final pair is not as bad as allowing clash to occur is left to future investigation.

4.2.3 Predicting order: Hiatus subset

The same mixed-effects logistic regression model presented in 4.2.1 was run on a subset of the flexible adjective data that excluded all pairs were hiatus was impossible (N=1766; Core Team et al., 2013).

Predicting: Order			
Fixed effect	Coefficient	SD	p-value
Clash Factor	3.75	1.68	p = 0.03(*)
Hiatus Factor	-0.40	0.16	p = 0.02(*)
Weight Factor	0.24	0.18	p = 0.18

Table 15: Mixed-effects logistic regression results: Hiatus subset

The results of the model are interpreted as follows. Since so many models were run on

the data, a p-value around 0.02 is not considered to be highly significant; therefore, none of the factors here were strongly significant. For weight, this is the same result as in the full model, but clash and hiatus were significant. Clash is believed to be less significant here because there was not a single pair that was categorized as having clash or avoided clash, all were impossible for clash (with a couple unknown). Hiatus may be less significant here, even though it is the subset of the data without any impossible hiatus pairs, because not only is there a virtually even split between true hiatus and avoided hiatus in this set, within each of those sets the split between prenominal and postnominal is virtually down the middle as well. In terms of what this means for the full model, it may be that a significant negative coefficient was found for hiatus due to a large number of impossible-hiatus pairs occurring in prenominal order (58% of impossible-hiatus pairs are prenominal).

4.2.4   Predicting order: Weight subset

The same mixed-effects logistic regression model presented in 4.2.1 was run on a subset of the flexible adjective data that excluded all pairs were weight was equal between the noun and adjective (N= 3728; Core Team et al., 2013).

Predicting: Order			
Fixed effect	Coefficient	SD	p-value
Clash Factor	0.39	0.24	p = 0.11
Hiatus Factor	-0.58	0.13	p = 6.29e-06*
Weight Factor	0.11	0.10	p = 0.29

Table 16: Mixed-effects logistic regression results: Weight subset

The results of the model are interpreted as follows. Only hiatus was found to be a significant effect, meaning it helps predict the surface order; however, clash and weight are not significant effects. The coefficient for hiatus is negative, showing the same effect as was found in the full model: adjectives remain postnominal in order to avoid hiatus in prenominal position. The results in this subset of the data, however, were different from the full model for clash, which is no longer a significant effect. This could be because the number of true

clash pairs paled in comparison to avoided/impossible, as well as unknown (33 clash vs. 3695 avoided/impossible or unknown).

4.2.5   Predicting adjective flexibility

Using the `glm` function from the `lme4` package in R, a logistic regression without random effects was run to see if the flexibility of an adjective (i.e., flexible vs. stricly prenominal and strictly postnominal combined) could be predicted by any of the phonological effects: clash, hiatus and/or weight (Core Team et al., 2013). Unlike the regression model predicting order, there were no random effects. The coefficients, standard deviations, and p-values for the fixed effects are reported in the table below (N = 14498).

Predicting: Adjective flexibility			
Fixed effect	Coefficient	SD	p-value
Clash Factor	0.45	0.08	p = 1.65e-08*
Hiatus Factor	0.52	0.04	p < 2e-16*
Weight Factor	0.01	0.04	p = 0.88

Table 17: Logistic regression results

The model is interpreted as follows. Clash and hiatus were found to both be significant effects, meaning that clash and hiatus both help predict if an adjective is flexible; however, weight was not significant. The coefficient values for clash and hiatus are both positive, meaning that if clash or hiatus occurred in the pair, it was likely that the order of the pair was fixed, rather than flexible[6]. The results of this model for clash and hiatus confirmed the hypothesis that, if word order is flexible, phonologically marked phenomena will tend to be avoided. However, this effect was predicted to not necessarily be found for hiatus, given that it is unclear if hiatus is active in Italian phonology. So while this is a positive finding for the general hypothesis that phonological markedness avoidance effects can be found in syntax, it was unexpected for hiatus. Unfortunately, this was effect not found to be significant for weight, where it was predicted.

---

[6]Recall that adjective type is coded as 1 if flexible and 0 if strictly prenominal or postnominal. Clash and hiatus are coded as 1 if avoided or impossible, and 0 if they occurred in the corpus.

The outcome of the same logistic regression model, but on different subsets of the data are reported in the Appendix. This was done for the same reason as stated above for the mixed-effects logistic regression model predicting order: to make sure the driving datapoints were not simply a large set of pairs in which the phonological effect was impossible, due to the coding. In the subsets of the data for clash and hiatus, the effects of those factors remained consistent, again suggesting a broader-scale effect of these phonological factors in line with the general hypothesis, while weight continued not to follow the prediction set out at the beginning of this work.

4.3 Monte Carlo Procedures

4.3.1 Clash

The Monte Carlo Procedure for stress clash was conducted using all instances of flexible adjective pairs in the corpus within which clash was possible, meaning every pair that had true stress clash or avoided stress clash was used.

In the portion of the real data, where pairs are flexible and clash is possible – ignoring all other clash types – the split between true clash and avoided clash is: 43% true clash and 57% avoided clash. Recall Prediction 1 stated in Section 3: Given that stress clash is actively avoided in Italian phonology, where word order is flexible, it will also be avoided syntactically via word-order manipulation in {noun, adjective} pairs. Following this prediction, the expected result from the Monte Carlo Procedure is that a random reshuffling of the data will result in more instances of true clash than in the actual data, and less instances of avoided clash than in the actual data. This prediction is confirmed by the results in Table 18 along with the trend found in the corpus for flexible vs. fixed adjectives which showed that the rate of clash was lower when the adjective was able to change position relative to the noun in comparison to pairs in which the position is fixed.

To confirm that the results of the Monte Carlo Procedure presented in Table 18 were not actually due to a correlation of these pairs with preferences based on phonological weight,

Monte Carlo Procedure	
Monte Carlo Distribution	50.3% (SD 3.4%) clash
Real Distribution	43% clash
One-sample T-test	
Effect size	15.7
p-value	$p < 0.01^*$

Table 18: Results of the Monte Carlo Procedure for clash

the Monte Carlo Procedure was conducted again, this time on flexible pairs which had words of equal syllable lengths (and in which clash was possible). This weight-neutralized result is shown in Table 19.

Monte Carlo Procedure	
Monte Carlo Distribution	50.4% (SD 5.9%) clash
Real Distribution	43% clash
One-sample T-test	
Effect size	9.1
p-value	$p < 0.01^*$

Table 19: Results of the Monte Carlo Procedure for clash, controlling for weight

Though the effect size is smaller, the result remains the same: the Monte Carlo distribution of clash is greater than the distribution in the real data, even when controlling for weight.

4.3.2 Hiatus

The Monte Carlo Procedure for vowel hiatus was conducted using all instances of flexible adjective pairs in the corpus within which vowel hiatus was possible, meaning every pair that had true vowel hiatus or avoided vowel hiatus was used.

In the portion of the real data, where pairs are flexible and hiatus is possible – ignoring all other hiatus types – the split between true hiatus and avoided hiatus is 50/50. Recall Prediction 2 stated in Section 3: Given that vowel hiatus can be said to *not* be actively avoided in Italian phonology, even where word order is flexible, it will not be avoided syn-

tactically via word-order manipulation in noun, adjective pairs. Following this prediction, the expected result from the Monte Carlo Procedure is that a random reshuffling of the data will result in a distribution of vowel hiatus that is *not* significantly different from the trend found in the actual data. The results in Table 20 confirm this prediction.

Monte Carlo Procedure	
Monte Carlo Distribution	49.8% (SD 1.4%) hiatus
Real Distribution	51% hiatus
One-sample T-test	
Effect size	-1.4
p-value	> 0.01

Table 20: Results of the Monte Carlo Procedure for vowel hiatus

In the Monte Carlo distribution, 49.8% of the pairs showed true hiatus, while in the real distribution the proportion is 51%. A one-sample T-test of these two distributions show that they are not significantly different, as predicted.

4.3.3   Weight

The Monte Carlo Procedure for phonological weight was conducted using all instances of flexible adjective pairs in the corpus in which the words were not of equal length, meaning every pair that was light- or heavy-final was used.

In the portion of the real data, where pairs are flexible and have unequal weight – ignoring equally-weighted pairs – the split between light-final and heavy-final pairs is 31% light-final vs. 69% heavy-final. Recall Prediction 3 stated in Section 3: Given that syntactic structure is sensitive to the phonological weight of NPs elsewhere in Italian, it will also be sensitive to phonological weight in flexible {noun, adjective} pairs. Following this prediction, the expected result from the Monte Carlo Procedure is that a random reshuffling of the data will result in a distribution of light-final pairs that is significantly higher than the trend found in the actual data. The results in Table 21 confirm this.

In the Monte Carlo distribution, 50% of the pairs were light-final, while in the real

Monte Carlo Procedure	
Monte Carlo Distribution	50.0% (SD 0.9%)
Real Distribution	31% light-final
One-sample T-test	
Effect size	205.2
p-value	$< 0.01^*$

Table 21: Results of the Monte Carlo Procedure for relative heaviness

distribution the proportion is 31%. A one-sample T-test of these two distributions shows that they are significantly different, as predicted.

4.4 Noun-Adjective Neighbors

4.4.1 Stress Clash

The same clash conditions were also observed between the {noun, adjective} target and its non-target neighbor. These data were collected to see if the constituency between noun and adjective played a role in how intolerable clash is. Nespor and Vogel (1979) argued that clash effects are seen only between tightly-grouped constituents. While it is possible for one of the neighbors to be a determiner, that is not guaranteed; so it is generally predicted here that clash with a neighbor will be better tolerated (i.e., observed at a higher rate) than clash within the noun-adjective pair, as a neighbor could be a word belonging to various categories, grouped or not grouped with the NP.

In the sample sequence: A [B C] D, [B C] is the noun-adjective pair, and A and D are each neighbors to that pair. Clash conditions were tested between A and B, and between C and D, but they were also tested between A and C and B and D. If one of the three (i.e., A, B, or C; B, C, or D) were unknown, clash was categorized as unknown. If the neighbor clashed, or had unavoidable clash, with both targets, clash was categorized as unavoidable. If the neighbor clashed with its adjacent target, but not with the non-adjacent target, clash was categorized as clash. If the neighbor clashed with the non-adjacent target but not with the adjacent target, then clash was categorized as avoided clash. Otherwise,

clash was categorized as impossible. Results of this test are in Table 22, along with the clash results for flexible {noun, adjective} pairs repeated for convenience.

FLEXIBLE {NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	180	3%
Unavoidable	0	0%
Clash	80	1%
Avoided clash	108	2%
Impossible clash	5,168	93%
Total	5,536	
TARGET AND NEIGHBOR PAIRS	COUNT	PERCENTAGE
Unknown	17,842	61.5%
Unavoidable	1,215	4.2%
Clash	2842	9.8%
Avoided clash	667	2.3%
Impossible clash	6,430	22.2%
Total	28,996	

Table 22: Comparative results of pair vs. neighbor clash data

There was an increase in unavoidable clash, and this is largely due to a neighbor being a determiner (monosyllabic, and therefore categorized as stressed by the database). Unfortunately much of the neighbor data did not appear in the phonological database, so it is categorized as unknown. There is still evidence, however, that suggests in the neighbor data, the predicted trend is found: clash was much more likely to be tolerated between a noun or adjective and its neighbor, than within a flexible {noun, adjective} pair. While much of the data remain unknown, this is a promising trend in the predicted direction, suggesting that constituency plays a role in how well clash is tolerated.

4.4.2 Vowel Hiatus

The same neighbor analysis was done for the vowel hiatus data. The prediction for hiatus remains mixed, since its tolerance by Italian phonology is also mixed. Results in Table 23 show more avoided hiatus between target and neighbor, than within target, contributing to the varied results for vowel hiatus found in this work.

FLEXIBLE {NOUN, ADJECTIVE} PAIRS	COUNT	PERCENTAGE
Unknown	434	8%
Unavoidable	182	3%
Hiatus	794	14%
Avoided hiatus	790	14%
Impossible hiatus	3,336	60%
Total	5,536	
TARGET AND NEIGHBOR PAIRS	COUNT	PERCENTAGE
Unknown	16,441	57%
Unavoidable	175	<0.1%
Hiatus	1,203	4%
Avoided hiatus	5,407	19%
Impossible hiatus	5,770	20%
Total	28,996	

Table 23: Comparative results of pair vs. neighbor hiatus data

While the avoidable hiatus results in the neighbors may seem like the opposite of what was expected (relatively more avoided hiatus pairs than hiatus pairs in neighbors vs. flexible {noun, adjective} pairs), this is largely due to vowel-final determiners. Unfortunately a lot of the data are unknown here as well. Given that in the flexible {noun, adjective} data, hiatus vs. avoided hiatus were about the same, or 50/50, and that the trend of more avoided hiatus in the neighbors being larger due to vowel-final determiners, these data are relatively uninteresting.

Neighbor data were not collected to look at relative heaviness as this phenomenon must occur within/between constituents.

4.5 Summary

Monte Carlo results indicated that clash occurred at higher rates in the reshuffled distribution, confirming my prediction; this was also true for weight, which is what was predicted, but is incongruous with the results from the regression models. Monte Carlo results for hiatus were insignificant, which follows from what was predicted.

Results of the mixed-effects logistic regression model predicting order from the three

phonological effects were consistent with the predictions for clash and hiatus, but not for weight. Clash was consistently found to push the order from the default postnominal to prenominal, in order to avoid it. Hiatus was initially found, in the full model, to cause pairs to remain postnominal in order to avoid hiatus in prenominal order; however, the subset of the data in which impossible hiatus was not included showed hiatus to no longer be a significant factor. Upon closer inspection, it was found that the significant effect of hiatus in the full model was caused by a large portion of the impossible-hiatus pairs occurring in prenominal order. This lack of a result for hiatus confirms the prediction that hiatus would likely not have an effect on syntactic order.

Results of the logistic regression model predicting adjective type show that avoiding clash and hiatus is correlated with adjective flexibility, but again weight was insignificant.

The implications of these results are discussed in the next section.

5   Discussion

5.1   Stress Clash

Results for clash were the strongest and most consistent out of the three effects investigated here, in support of my hypothesis. The mixed-effects logistic regression found that clash was able to predict order, in the direction that clash occurrence in postnominal/default order pushed the pair to a prenominal order in which clash was avoided. This result held even in the smaller subset of data in which impossible clash pairs (coded alongside avoided clash) were excluded.

This is consistent with the prediction that a more marked syntactic order is tolerated if it means that a phonologically-marked phenomenon like clash, which is active in Italian phonology, is avoided. This suggests an effect of phonology on syntax via word-order in these {noun, adjective} pairs.

Additionally, the logistic regression predicting adjective flexibility showed that avoided

clash was consistent with an adjective being flexible, rather than strictly ordered. This remained the result in the smaller model (see Appendix). This also confirms the prediction that flexibility in word order will allow for phonologically-marked phenomena to be avoided at a higher rate.

This was also confirmed by the results of the Monte Carlo Procedure for clash, in which the randomly reshuffled distribution showed a higher rate for clash/lower rate for avoided clash than the actual corpus data. This means that the rate of avoided clash found in the real data is not likely to be random, but rather is hypothesized to be the result of word re-ordering.

Finally, in the {noun, adjective} neighbor results, clash was found to occur at higher rates between a member of the pair and its neighbor, which did not necessarily form a syntactic constituent, unlike {noun, adjective} which forms an NP. This confirms a claim from the literature asserting that phonological effects on syntax are stronger when they occur between words or phrases that are organized in some way syntactically.

Taken together, these results strongly support the claim that there is an effect of phonology on syntax in these data for Italian.

5.2   Vowel Hiatus

Results for hiatus were fairly varied, in line with my hypothesis that, since hiatus is not very active in Italian phonology, it will not have strong effects on syntax. The mixed-effects logistic regression found that hiatus helped predict order, in that hiatus caused pairs to remain postnominal in order to avoid hiatus in prenominal position; however, in the smaller model, this effect was not present and upon closer inspection of the data, the original effect could be attributed to a large portion of impossible-clash pairs occurring in prenominal order.

The logistic regression predicting flexibility did show that avoided hiatus was consistent with an adjective being flexible, and while this remained true in the smaller model (so it could not be attributed to a trend found in the impossible-hiatus pairs, like the model predicting

36

order; see Appendix), it should be regarded with suspicion since it is inconsistent with the regression models predicting order and results of the Monte Carlo Procedure.

The Monte Carlo distribution for hiatus was not significantly different from the trend found in the actual data, indicating that hiatus occurrence/avoidance is near-chance.

Taken together, these results do not show strong evidence that hiatus has an effect on the word order of {noun, adjective} pairs in Italian, but it was predicted to not necessarily have one. In Italian, hiatus is tolerated at different levels depending on the vowel qualities involved. A high vowel preceding a non-high vowel often becomes a glide (Kramer, 2009). If the data were narrowed down to pairs which involved a high vowel, the results might then show some sort of avoidance via word order. This deeper investigation of hiatus is left to future work.

5.3 Phonological Weight

Results for weight were unexpected as they are inconsistent with the hypothesis. Given that weight can be said to be active in Italian (see 2.4.3 for data on OVS word order restrained by phonological weight), it was predicted to be avoided by word reordering in the {noun, adjective} pairs.

Mixed-effects logistic regression results predicting order, however, showed no significant effect of weight, even when equal-weight pairs were excluded. In addition, logistic regression results predicting adjective flexibility also showed no significant effect of weight, and were not in line with the prediction even in the smaller model excluding equal-weight pairs (see Appendix).

Only the results from the Monte Carlo Procedure were in line with predictions for weight, with significantly more pairs in the reshuffled distribution being light-final than in the actual data.

The inconsistency between the results from the regression models and those from the Monte Carlo Procedure may be explained by the following: in comparing weight data for

prenominal vs. postnominal pairs (among flexible adjectives), it was found that the proportion of light-final and equal-weight pairs is similar in both sets: 51% of pairs with prenominal order are light-final or equal-weight compared to 57% of pairs with postnominal order. It is possible that these distributions were too similar for the model to be able to reliably correlate light-final/equal-weight with one order over the other; thus, the insignificant coefficient for the weight factor in the model predicting order of flexible adjectives.

In comparing weight data for flexible vs. inflexible pairs, it was found that the proportion of light-final pairs (excluding equal-weight pairs) is virtually the same in both sets: 30% of flexible pairs are light-final and 29% of inflexible pairs are light-final. This explains the lack of result for the logistic regression model predicting adjective flexibility, because the trends across the two adjective types may also be too similar.

The result from the Monte Carlo distribution is inconsistent with the regression model results because while it is possible that in the flexible dataset, light-final pairs occur less often than if words were more randomly combined, it could also be that this is true regardless of adjective type.

As for consistency across adjective types, this could have arisen from a diachronic process that preferred longer strictly-postnominal adjectives and shorter strictly-prenominal adjectives so as to increase chances for a heavy-final pair. As for word order, another avoidance strategy may be at play, such as synonym selection. Rather than reordering, speakers select a synonym for the adjective or noun that allows for a heavy-final pair. So while the results shown here appear inconsistent or against the original hypothesis, it is possible that the avoidance of light-final constituents is manifested in other ways, via language change, lexical selection, or other strategies. A deeper exploration of these alternative processes is left to future work.

5.4 Neighbors

The different types of clash as well as vowel hiatus were also investigated between each member of the {noun, adjective} pairs and the word occurring to the left and right of the pair. This was done in order to see if constituency played a role in the relative intolerance of marked phonological phenomena, as has been proposed by Nespor and Vogel (1979). While much of the data are unfortunately unknown[7], there is a promising trend in the comparison of the clash data for the two different syntactic environments. Within flexible {noun, adjective} pairs, clash was avoided more than it was allowed; and this was the opposite among the neighbor pairs. While more data would certainly allow for a better claim to be made, this does suggest that clash is less tolerated between words that form a close constituent. This means that for clash, not only is there an influence of phonological markedness on syntactic structure, but the reverse is true as well.

5.5 A Note on Semantics

As previously stated, the semantics of the {noun, adjective} pairs analyzed here was not the focus of this work. While the order of {noun, adjective} can have an effect on the meaning of this NP, this is not always the case and this effect is largely left to future work. A preliminary look, however, was taken at the clash vs. avoided clash pairs in flexible adjectives. All unique pairs were presented to a linguistically-aware native speaker in both orders: the order in which they were found in the corpus, and the reverse. The speaker was asked to judge if they perceived a difference in meaning between the pairs, and which pair, if either, they preferred. The full data of this informal experiment can be found in the Appendix, but there are some results worth highlighting here.

Many of the avoided clash pairs had differences in meaning. Sometimes the speaker noted why they preferred one order over the other for these pairs which had a meaning change and two main reasons cited were that one order was a fixed expression (and the

---

[7]Due to many words from the corpus that do not appear in the PhonItalia lexical database.

other ungrammatical), and that one order had a more "literary" reading. Pairs marked as literary were those with a normally prenominal (but still flexible) adjective in the postnominal position. One such is example is *ultima attività* 'last activity' (typical order for this adjective) vs. *attività ultima* (atypical order for this adjective, but general default order for {noun, adjective pairs}), wherein *attività ultima* was marked as literary.

Some adjectives changed readings at a semantic, rather than more pragmatic, level. This was most noted by the speaker with the adjective *stesso* 'same.' In prenominal position, this adjective has the typically glossed meaning 'same,' but in postnominal position, the speaker noted it meant something more like 'itself.' An example from the data is shown below[8].

(14) a. *stesse*   *società*   PRENOMINAL (attested)
          same.FEM.PL   societies
          'same societies'

      b. *società*   *stesse*   POSTNOMINAL (created)
          societies   same.FEM.PL
          'societies themselves'

While there were a few other pairs in the avoided clash category with noted semantic differences based on order (see Appendix), it is unclear how pervasive this is, and how large of an effect the semantics of these pairs has on the surface order of {noun, adjective}.

In comparison, the plurality of pairs from the corpus where clash did occur were categorized by the speaker as fixed expressions, which was not true of the avoided clash pairs. While the semantics of these pairs in general needs further investigation, the contrast between the avoided clash and true clash pairs in terms of the semantic makeup of the adjectives or pairs involved suggests that clash is more likely to be tolerated if there is a strong semantic reading associated with that particular word order. Over time, if the semantic relationship with word order is prioritized over the phonological one, this could lead to the crystallization of forms such as those seen in the clash data: only one order is grammatical, associated with a particular meaning, and clash is tolerated.

---

[8]The prenominal, avoided class pair in (14a) was what was found in the corpus. The reverse pair in (14b) was created for the informal experiment. Both are grammatical.

5.6 Implications for a Syntactic Model

It is clear from these data that word order is used as a strategy to avoid stress clash, a phenomenon active in Italian phonology. Though results were not compelling for hiatus or phonological weight, by taking a closer look, we may find that these phonologically-marked phenomena are avoided as well (in the case of hiatus), or by alternative methods (e.g., lexical selection, in the case of weight). These processes in general may be part of a broader guiding principle in linguistics: euphony[9]. This principle, while grounded in phonology, may exist at a higher cognitive level, affecting all components of language: semantic, syntactic, phonological and phonetic.

Euphony, however, may often be at odds with the speaker's need to effectively communicate her message. So while clash was found to affect word-order, though there's evidence elsewhere in the language for its effects, phonological weight did not. As discussed in 5.4, this result could be obscured by the effects of alternative processes like language change and/or lexical selection, but it could also be the result of an outranking of more precise semantics over more euphonious phrasing.

For example, some of the flexible pairs that remained light-final – despite the structural ability to move – seemed to be restricted by a change in meaning that would have resulted from the pair occurring in the alternative order. A couple of pairs demonstrating this are shown below.

(15) a. *natura*   *stessa*   POSTNOMINAL (attested)
          nature   same.FEM.PL
          'nature itself'
      b. *stessa*   *natura*   PRENOMINAL (change in meaning)
          nature   same.FEM.PL
          'same nature'

Though the prenominal order is possible for both of these adjectives and would render the phrase more euphonious, with the heavier noun in phrase-final position, this would change

---

[9]Language that is perceived as rhythmic or harmonious in some way.

(16) a. *femmina*   *buona*   POSTNOMINAL (attested; ambiguous)
          woman   good.FEM.SG
          'good woman' **or** 'good-hearted woman' *non-intersective* or *intersective*
     b. *buona*   *femmina*   PRENOMINAL (disambiguated)
          woman   good.FEM.SG
          'good woman'   *non-intersective*

the meaning or sets of meanings available for the phrase and greatly, if not fatally, put the speaker at risk of being misunderstood.

The percentage of light-final flexible-adjective phrases, while not the plurality, made up a non-trivial proportion of the data at 23%. This semantic motivation is hypothesized to be at the heart of this trend, and is left to future investigation.

The interaction of phonology and syntax evidenced in these data, at least for clash, could have arisen from diachronic processes, or it could be the result of considerations of these sometimes competing factors (semantic precision and euphony) made on-line. If diachronic, this would be theoretical evidence for a more general linguistic preference for euphony. If on-line, this would be evidence for something like a cyclic relationship between syntax, semantics, and phonology, where a general idea is formed, then a cycle of constituent construction with verification by semantics and phonology occurs before an optimal or sufficient phrase is chosen and uttered. A diachronic investigation of these trends analyzed here, as well as experimental work more directly observing what speakers are doing, are both needed to better elucidate this finding.

6   Conclusion and Future Directions

The hypothesis for this work, that actively phonologically-marked phenomena in a language are also avoided syntactically, was found to be true for at least one phenomenon investigated here: stress clash. Clash was avoided more often when adjectives had flexible order, and it was shown that prenominal order correlated with avoided clash, suggesting that if clash occurs in postnominal, default order, the order will be pushed to prenominal to avoid clash.

Clash was found to occur at higher rates if pairs were more randomly constructed (i.e., in the Monte Carlo distribution), and if the two words did not necessarily belong to the same constituent (i.e., in the neighbor results). Taken together, these results present a strong argument for phonological effects on syntax at the interface.

Results for hiatus and weight were not as compelling, but further investigation of these phenomena is needed. Future work will take a closer look at the effect of high-vowels in the hiatus data, and potential evidence of lexical selection and/or diachronic processes as explanations for the results found for weight. Semantic effects need also to be more carefully considered. More corpus work looking at older stages of Italian or more conservative dialects, as well as experimental work looking more directly about what speakers are doing when constructing these phrases would both contribute greatly to the investigation of this topic.

Additionally, an expanded look at phonological markedness effects on syntax including other syntactic structures and phonological phenomena in additional languages would add to a growing body of work showing these effects at the syntax-phonology interface.

References

Borrelli, D. A. (2013). *Raddoppiamento sintattico in Italian: a synchronic and diachronic cross-dialectical study.* Routledge.

Breiss, C. and Hayes, B. (2019). Phonological markedness effects in sentence formation. *Ms., University of California, Los Angeles.*

Cardinaletti, A. (2010). On a (wh-) moved topic in italian, compared to germanic. *Advances in comparative Germanic syntax*, pages 3–40.

Chomsky, N. (2010). Lectures on government and binding: The pisa lectures. In *Lectures on Government and Binding.* De Gruyter Mouton.

Cinque, G. (2010). *The syntax of adjectives: A comparative study*, volume 57. MIT press.

Core Team, R. et al. (2013). R: a language and environment for statistical computing. *R Foundation for statistical computing, Vienna.*

Frascarelli, M. (2000). *The syntax-phonology interface in focus and topic constructions in Italian*, volume 50. Springer Science & Business Media.

Goslin, J., Galluzzi, C., and Romani, C. (2013). Phonitalia: a phonological lexicon for Italian. *Behavior Research Methods.*

Inkelas, S. and Zec, D. (1990). *The phonology-syntax connection.* University of Chicago Press.

Inkelas, S. and Zec, D. (1995). Syntax-phonology interface. the handbook of phonological theory, edited by john goldsmith, 535-549.

Jun, S.-A. and Fougeron, C. (2000). A phonological model of french intonation. In *Intonation*, pages 209–242. Springer.

Kramer, M. (2009). *The phonology of Italian.* Oxford University Press on Demand.

Liberman, M. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic inquiry*, 8(2):249–336.

Martin, A. (2011). Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language*, 87(4):751–770.

Nespor, M. and Vogel, I. (1979). Clash avoidance in italian. *Linguistic Inquiry*, 10(3):467–482.

Nespor, M. and Vogel, I. (1986). *Prosodic Phonology.* Dordrecht: Foris.

Shih, S. S. and Zuraw, K. (2017). Phonological conditions on variable adjective and noun word order in tagalog. *Language*, 93(4):e317–e352.

Varis, E. E. (2012). *The Spanish feminine el at the syntax-phonology interface.* University of Southern California.

Appendix

Smaller logistic regression models predicting adjective flexibility

Clash subset

The same logistic regression model from 4.1 was also run on a subset of the data, excluding pairs with impossible clash to predict adjective flexibility. Results are reported in Table 24 (N = 1281).

Predicting: Adjective flexibility			
Fixed effect	Coefficient	SD	p-value
Clash Factor	1.76	0.18	p < 2e-16*
Hiatus Factor	0.30	0.19	p = 0.11
Weight Factor	0.28	0.13	p = 0.03(*)

Table 24: Logistic regression results: Clash subset

Clash remained significant with a positive coefficient, indicating that clash avoidance was associated with adjective flexibility, confirming my prediction. Now also significant is the weight factor, indicating that, within this subset of the data, heavy-final pairs helped predict that the adjective was flexible. While this follows my prediction, this effect was not significant in the full model. Hiatus is no longer significant in this smaller model.

Hiatus subset

The same logistic regression model from 4.1 was also run on the subset of the data excluding pairs with impossible hiatus. Results are reported in Table 25 (N = 5391).

Predicting: Adjective flexibility			
Fixed effect	Coefficient	SD	p-value
Clash Factor	-0.88	0.45	p = 0.05
Hiatus Factor	0.28	0.06	p = 3.04e-06*
Weight Factor	-0.03	0.07	p = 0.64

Table 25: Logistic regression results: Hiatus subset

Hiatus remained significant, confirming the prediction that hiatus avoidance was associated with adjective flexibility. Clash is no longer significant in this smaller model. Weight remained an insignificant factor as well.

Weight subset

The same logistic regression model from 4.1 was also run on the subset of the data excluding pairs that were equal weight to predict adjective flexibility. Results are reported in Table 26 (N = 10483).

Predicting: Adjective flexibility			
Fixed effect	Coefficient	SD	p-value
Clash Factor	0.69	0.09	p = 6.65e-15*
Hiatus Factor	0.33	0.05	p = 1.43e-12*
Weight Factor	-0.15	0.05	p = 0.001*

Table 26: Logistic regression results: Weight subset

In this smaller model, clash and hiatus remained significant factors with positive coefficients, which is in line with my predictions. However, weight is also significant in this model (where it is not in the full model) with a negative coefficient. This means that light-final pairs were actually associated with flexible adjectives; this is against the prediction.

	PAIO 1			PAIO 2		ti sembri differenti?	quale preferisci?
	TIPO 2		**CLASH PAIRS: ACTUAL DATA = PAIO 1**				
1	martedì	scorso	vs.	scorso	martedì	Y	1
2	città	vecchia	vs.	vecchia	città	N	2
3	età	media	vs.	media	età	1 = fixed expression	
4	maggior	peso	vs.	peso	maggiore	1 = fixed expression	
5	onestà	tipica	vs.	tipica	onestà	2 = fixed expression	
6	maggior	grado	vs.	grado	maggiore	need context	
7	necessità	pubbliche	vs.	pubbliche	necessità	need context	
8	sanità	pubblica	vs.	pubblica	sanità	1 = fixed expression	
9	indennità	pari	vs.	pari	indennità	Y	2, 1 = *
10	incolumità	pubblica	vs.	pubblica	incolumità	Y	2
11	maggior	termine	vs.	termine	maggiore	need context	
12	contabilità	pubblica	vs.	pubblica	contabilità	Y	2
13	maggior	parte	vs.	parte	maggior	1 = fixed expression	
14	autorità	pubbliche	vs.	pubbliche	autorità	N	either
15	maggior	numero	vs.	numero	maggiore	1 = fixed expression	
16	età	minima	vs.	minima	età	1 = fixed expression	
17	specialità	tipiche	vs.	tipiche	specialità	1 = fixed expression	
18	città	vecchia	vs.	vecchia	città	Y	context
19	autorità	pubbliche	vs.	pubbliche	autorità	need context	
20	città	sacra	vs.	sacra	città	1 = fixed expression, but both work	
21	città	santa	vs.	santa	città	1 = fixed expression, but both work	
22	età	minima	vs.	minima	età	1 = fixed expression	
23	giovedì	prossimo	vs.	prossimo	giovedì	1 = fixed expression	
24	maggior	uso	vs.	uso	maggior	need context	
25	continua	fonte	vs.	fonte	continua	1 = fixed expression	
26	maggior	tempo	vs.	tempo	maggiore	1 = fixed expression	
27	città	medie	vs.	medie	città	need context	
28	città	grandi	vs.	grandi	città	1 = fixed expression	
29	film	turco	vs.	turco	film	Y	1, 2 = *
30	eredità	stessa	vs.	stessa	eredità	Y	2, 1 = *
31	responsabilità	proprie	vs.	proprie	responsabilità	Y	2, 1 = *

Notes:
1. Some pairs from the corpus were judged ungrammatical (30 and 31).
2. Many pairs were fixed expressions, and many others needed context for the participant to make judgments

AVOIDED CLASH PAIRS, ACTUAL DATA = PAIO 1

	PAIO 1			PAIO 2		ti sembri differenti?	quale preferisci?
1	simile	gravità	vs.	gravità	simile	N	N/A
2	giovani	sub	vs.	sub	giovani	Y	1
3	grandi	capacità	vs.	capacità	grandi	N	1
4	grande	facilità	vs.	facilità	grande	Y	1
5	stesse	società	vs.	società	stesse	Y	context
6	modulo	continuo	vs.	continuo	modulo	N	N/A
7	visite	continue	vs.	continue	visite	N	N/A
8	pubblica	utilità	vs.	utilità	pubblica	N	N/A
9	giusta	indennità	vs.	indennità	giusta	Y	N/A
10	minima	unità	vs.	unità	minima	Y	2, 2 = fixed
11	ultima	volontà	vs.	volontà	ultima	Y	1, 2 = literary
12	pubblica	autorità	vs.	autorità	pubblica	Y	1, 2 is also fine
13	sola	metà	vs.	metà	sola	Y	1, 2 = ?
14	acqua	continua	vs.	continua	acqua	N	N/A
15	pari	dignità	vs.	dignità	pari	Y	1, 1 = fixed
16	piena	libertà	vs.	libertà	piena	Y	1, 2 = literary
17	pari	opportunità	vs.	opportunità	pari	Y	1 = fixed, 2 = *
18	piena	parità	vs.	parità	piena	Y	1, 2 = ?
19	stesse	modalità	vs.	modalità	stesse	Y	context
20	altre	attività	vs.	attività	altre	Y	1, 2 = ?
21	altre	città	vs.	città	altre	Y	1
22	primo	gol	vs.	gol	primo	Y	1
23	alta	qualità	vs.	qualità	alta	Y	1, 1 = fixed
24	grandi	quantità	vs.	quantità	grandi	Y	1
25	altre	società	vs.	società	altre	Y	1
26	prima	attività	vs.	attività	prima	Y	1
27	ultima	attività	vs.	attività	ultima	Y	1, 2 = literary
28	primo	film	vs.	film	primo	Y	1, 2 = ?
29	primo	post	vs.	post	primo	Y	1, 2 = ?
30	certa	tempestività	vs.	tempestività	certa	Y	context
31	stesse	difficoltà	vs.	difficoltà	stesse	Y	context
32	nuove	opportunità	vs.	opportunità	nuove	Y	1
33	giovane	età	vs.	età	giovane	Y	1, 2 = ?
34	pari	opportunità	vs.	opportunità	pari	Y	
35	grande	opportunità	vs.	opportunità	grande	Y	different uses
36	bassa	produttività	vs.	produttività	bassa	N	1
37	alta	produttività	vs.	produttività	alta	N	1
38	nuove	unità	vs.	unità	nuove	N	slight change
39	grandi	città	vs.	città	grandi	N	1

40	grandi	proprietà	vs.	proprietà	grandi	N	1
41	grandi	difficoltà	vs.	difficoltà	grandi	N	1
42	grande	velocità	vs.	velocità	grande	N	1, 2 = ?
43	nuovo	film	vs.	film	nuovo	N	
44	prima	metà	vs.	metà	prima	N	1, 2 = ?
45	certa	regolarità	vs.	regolarità	certa	Y	
46	pubblica	utilità	vs.	utilità	pubblica	N	
47	alta	densità	vs.	densità	alta	Y	1, 1 = fixed
48	nuovo	film	vs.	film	nuovo	N	1
49	grandi	novità	vs.	novità	grandi	N	1
50	grande	varietà	vs.	varietà	grande	N	1
51	buona	qualità	vs.	qualità	buona	Y	1, 1 = fixed

Notes
1. Pairs with stesso/stessa/stessi/stesse have a change in meaning: prenominal = 'same X', while postnominal = 'X themselves'
2. Pair with certa (30) has a change in meaning: prenominal = 'some X, while postnominal = 'safe/secure X'
3. Pair with nuove (38) has a change in meaning: prenominal = 'new unit (of monuments)', while postnominal = 'new group (of people)'
4. Pairs with alto/alta/alti/alte often have a change in meaning or are fixed expressions.
Participant noted *città alta* = 'uptown', *alto livello* = 'next level (e.g., video game)', *livello alto* = 'high quality'; and in general the prenominal order is perceived with this adjective as fixed/bookish