Tel-Aviv University
The Lester and Sally Entin Faculty of Humanities
Department of Linguistics

Abbreviations:
A Unified Analysis of Acronym Words, Clippings, Clipped Compounds, and Hypocoristics

Thesis submitted for the M.A Degree at Tel-Aviv University

by:
Gila Zadok

Thesis Advisor:
Dr. Outi Bat-El

October, 2002

Acknowledgements

Although writing is a painstaking solitary task, this work would not have been possible if it were not for the guidance and support of the people surrounding me. First and foremost, I would like to extend my deepest gratitude to my thesis advisor, Outi Bat-El. Outi has always supported me in everything I did and encouraged me to do things I would never have attempted of my own accord. She invested much time and effort in me, and did so with endless patience. And despite the hard work and the insecurities that go with it, she made this entire endeavour fun. I could not have wished for a better and more attentive advisor.

I feel extremely lucky to have enjoyed my studies as much as I have. I believe that this is due primarily to Avivit Ben-David and Galit Adam, who from the very beginning made me feel welcome in what should have been a threatening environment. I thank them for their persistent interest in my progress and for being good listeners and friends. I thank Charles Kisseberth, from whom I have learned so much; it has truly been a privilege.

I owe very special thanks to Galit Adam for her valuable insights and suggestions and indeed for all her help. Galit has generously spent time reading my paper, commenting on it and answering all my questions. I have benefited immensely from her acquaintance both academically and personally.

Thanks to my friends and colleagues at BVR Systems for hearing my talk and for undertaking to collect some of the Hebrew data. Specifically, I'd like to mention Gilad Yavetz, with whom I made a pact in this respect, and Yariv Gershony, for many years of

friendship and for being the most colourful person I have ever had the privilege of knowing.

Looking back, I realise that my entire academic experience was influenced by some major events and changes. I thank my dear friends Yafit and Ronit for sticking by me through thick and thin. I thank my mother, Jasna, my brother, Uri, and his wife Sharon, for caring so much and for sharing these events with me. I dedicate this work to my father, who did not get to know me as I am today.

Table of Contents

Chapter 1. Introduction

1.1. Abbreviations

Abbreviations, acronyms, initialisms, clippings, reduced compounds, are all terms found in various studies and dictionaries in their attempt to classify the various types of shortened forms. These forms have many shapes which make their classification difficult. Scholars who attempted to classify these terms only added to the confusion. Kennedy (1935) views words such as *Arco* 'American Radio Co.' as blends; Bauer (1983) sees words as *stagflation* (from *stag*nation + in*flation*) as blends, but *Arco*, along with forms such as WASP, where the initial letters of every word in the base form a pronounced word, as acronyms; Gales' *Acronyms Dictionary* classifies the blends *motel*, *brunch*, and *smog* as acronyms; a search on the World Wide Web reveals that all initialisms are listed as acronyms, even when the items are unpronounceable; and today, *abbreviation* denotes all kinds of shortenings, rather than only what Cannon (1989) terms "legitimate abbreviations", such as FBI and CIA.

Classification is difficult even when definitions exist, because some of the definitions overlap, and because some of the items combine what are traditionally considered separate processes. Cannon's attempt to distinguish between clippings and acronyms illustrates this point. "An acronym preserves only the initial part(s) of a multiword source" (p.108). According to this definition, *sci-fi* 'science fiction' is an acronym, but *Cogas* 'coal, oil, gas' is not, as *gas* is not reduced. However, border cases such as *Autoland* 'automatic landing' raise the question of "how many letters/sounds/syllables must be lost (or preserved) before an item can be classed as an acronym" (p.108).

In this study, for descriptive purposes, I will refer to four types of abbreviations:

a. Acronym Words - words comprised of only one segment from each base word.

b. Clipped Compounds - words comprised of more than one segment from at least one base word.

c. Clippings - words that are formed from a single base word.

d. Hypocoristics - abbreviated names (i.e. clippings of personal names).

Below are some examples of these abbreviations from the three different languages to be considered in this study.

(1) Acronym Words (AW)

Base	AW	Gloss
English		
American **S**tandard **C**ode for **I**nformation **I**nterchange	ASCII (ǽski:)	
North **A**tlantic **T**reaty **O**rganisation	NATO (néɪtəʊ)	
Hebrew		
ʔatomi **b**iyologi **x**imi atomic biological chemical	ʔábax	'atomic biological & chemical' (warfare)
cva **h**agana **l**e-israel army defence to-israel	cáhal	'Israeli Defence Force'
Serbo-Croatian		
Savezno **I**zvršno **V**eće	sív	'Federative Executive Assembly'
Akademsko **K**ulturno **U**metnićko **D**ruštvo	ákud	'Academic cultural artistic company'

(2) Clippings

Base	Clipping	Gloss
English		
memorándum	mémo	
advertísement	ád	
Hebrew		
súpermarket	súper	‘supermarket’
trigonoométria	trígo	‘trigonometry’

(3) Clipped Compounds (CC)

Base	CC	Gloss
English		
situation **com**edy	sitcom	
formula **tran**slation	fortran	
Hebrew		
menahel **kl**ali	mankál	‘general manager’
ʔanašim **xa**šuvim **m**eod	ʔaxám	very important people’ (VIP)
Serbo-Croatian		
ekoloških **o**rganizaciya **s**rbiye	ékos	‘Ecology organisation of Serbia’
komunalno **gra**jevinsko **p**reduzeće	kómgrap	‘Community building company’

(4) Hypocoristics

Base	With Suffix	Without Suffix
English		
samántha	sámi	sám
wílliam	wílli	wíll
Hebrew		
daniél	dáni	
revitál	révi	
Serbo-Croatian		
yélena	yéla	
bórislav	bóra	

Former analyses of these types of abbreviation processes can be found in the linguistic literature. For example, Bat-El (1994) for Hebrew acronyms, Itô and Mester (1997) for German hypocoristics, Mester (1990) and Poser (1990) for Japanese, Piñeros (1999) for Spanish, and Weeda (1992) for French and English. However, these studies usually treat each phenomenon separately and do not attempt to find some common grounds to them or to discover the similarities between the languages. If these are indeed independent processes, one could plausibly expect to find two different forms derived from the same base: an AW and a Clipped Compound, for example, as is the case, for example, with affixation, where different affixes are attached to the same base forming two different (though related) words (e.g. *clearly* and *clearness*). However, we do not find two independent organisations with different names abbreviated from the same base, e.g. mothers against gambling

> mag/motag, or computer aided design > cad, compad. That this does not occur suggests that they are all subject to the same system.

Abbreviations are often considered to be marginal and extragrammatical (Dressler & Karpf 1994) as they are consciously, and therefore unnaturally, coined, and even chaotic in terms of the deleted segments (Bauer 1988). If indeed they are marginal, extragrammatical, or chaotic, they are undeserving of any grammatical treatment. I will argue, however, following Bat-El (2000a), that the principles governing abbreviations are drawn from those responsible for the grammar of natural languages.

In this study I will describe not only the similarities between languages with respect to these types of abbreviations, but also their common properties. I will argue that the four types of abbreviations under consideration follow the same principles, and that the differences between them are limited and systematic.

1.2. Theoretical Background

The chaotic nature of abbreviations (Bauer 1988), refers to the number of deleted segments. This number differs from one abbreviation to another (e.g. *exam<ination>* - 7 deleted segments, *ad<vertisement>* - 10 deleted segments, *gym<naesium>* - 6 deleted segments). Within a rule based approach, the target must be specified, either in phonological terms (syllables, segments), or in morphological terms (morphemes). However, the target differs from one form to another, as the number of remaining segments varies (in the above examples 4, 2, 3 respectively). Also, neither the deleted segments nor the remaining ones necessarily form a morphological unit. Therefore, no rule (phonological or morphological) can be stipulated to generate these abbreviations.

In this study I will claim that if we look at the output syllables, abbreviation processes are not at all chaotic. The theoretical framework that allows direct reference to the output is Optimality Theory (Prince and Smolensky 1993). According to this theory, a generator provides several candidates for each input. These candidates are evaluated by the language's constraint hierarchy. The constraints are universal, and may be violated in order to satisfy a higher ranked constraint. The optimal candidate, the one to emerge as the surface form, is the one which least violates the constraint hierarchy, i.e. the one which best satisfies the higher ranked constraints and minimally violates the lower ranked constraints.

There are two types of universal constraints: markedness and faithfulness. Markedness refers to the properties of the output which are shared by all languages and to the possible range of variation between them. For example, the markedness constraint ONSET, which requires that syllables have onsets, is based on the typological observation that languages that have onsetless syllables also have syllables with onsets, but not vice versa. Syllables with onsets are therefore unmarked with respect to onsetless ones. Hence, unmarked properties are shared by all languages and languages differ in the range of marked properties they display. Faithfulness constraints refer to the relations between two forms, and are at the heart of correspondence theory (McCarthy and Prince 1995). They require identity between corresponding forms, whether they be underlying and surface (input-output), two surface forms (output-output), base and reduplicant, etc.

The properties shared by abbreviated forms can be defined in prosodic terms, a tenet of the theory of Prosodic Morphology. The goal of Prosodic Morphology (McCarthy and Prince 1986 et seq.) is to derive morphological regularities from

general properties of morphology, phonology, and the interface between morphology and phonology.

Following Selkirk (1980a,b) the prosodic units are organised in a hierarchy (5), which dictates that a prosodic word minimally contain a foot.

(5) The Prosodic Hierarchy (Selkirk 1980a, 1980b)

PrWd	(Prosodic Word)
|	
Ft	(Metrical Foot)
|	
σ	(Syllable)
|	
μ	(Mora)

Metrical feet can be either syllabic or moraic. The inventory of possible feet is provided in (6), where L denotes light (monomoraic) syllables, and H denotes heavy (bimoraic) syllables; σ indicates a syllable of any type.

(6) Foot Inventory: (Hayes 1987, 1995; Kager 1989, 1991)

Iambic	Trochaic	Syllabic
LH, LL, H	H, LL	σσ

The constraint responsible for the absence of a single light syllable from the foot inventory is FTBIN (7), which requires feet to be binary either syllabically (two syllables), or moraically (two moras).

The prosodic hierarchy together with the foot binarity requirement derive the notion of "Minimal Word". The Minimal Word must contain at least a di-syllabic or bimoraic foot.

McCarthy and Prince (1994) propose that the Minimal Word emerges when three prosodic constraints are strictly respected: FTBIN, PARSEσ, and ALL-FT-L/R.

(7) FTBIN Feet are binary on a syllabic or moraic level.[1]

PARSEσ All syllables are parsed into feet (following the prosodic hierarchy)

ALL-FT-L/R Align (Ft, L/R; PrWd, L/R) The left/right edge of all feet coincide with the left/right edge of the prosodic word.

(This constraint penalises for every syllable intervening between the (right or left) edge of the foot and that of the prosodic word.)

The empirical support for the minimal word is drawn from languages which place prosodic restrictions on the minimal size of well-formed content words. In Yidin^{y} (Dixon 1977) and Dyribal (Dixon 1972) words are minimally disyllabic; Japanese derived words (Poser 1990) and Estonian content words (Prince 1980) must consist of minimally a bimoraic foot. As shown below, such restrictions may block or trigger phonological processes.

In Estonian, a final vowel is deleted in the nominative case (8a), unless the resulting form violates the minimality requirement (8b):

(8) Estonian (Prince 1980)

	Base	Nom. sg.		Gloss.
a.	tænava	tænav		'street'
	konna	kon:n		'frog'
	matsi	mat:s		'lout'
b.	kana	kana	*kan	'hen'
	koi	koi:	*ko	'clothes-moth'
	maa	maa:	*ma	'country'

In Lardil, final vowels are also deleted in the Nominative (9a), but are preserved so as not to violate the minimality requirement (9b). Moreover, to satisfy the minimality requirement, subminimal roots are augmented to two moras (9c):

[1] In quantity sensitive languages a binary foot can either be disyllabic or bimoraic. In quantity insensitive languages, this constraint can only be satisfied by a disyllabic foot.

(9) Lardil (McCarthy and Prince 1994)

	UR	Nominative		Gloss.
a.	/mayara/	mayar		'rainbow'
	/kantukantu/	kantukan[2]		'red'
b.	/wiʈe/	wiʈe	*wiʈ	'inside'
	/mela/	mela	*mel	'sea'
c.	/wik/	wika		'shade'
	/ter/	tera		'thigh'

Within Optimality Theory, constraints may be violated in order to satisfy a higher ranked constraint. Perfect satisfaction of the constraints in (7) is possible, and would result in a single disyllabic foot (or bi-moraic in quantity sensitive languages). Satisfaction of these constraints may, however, entail loss of material and violation of lower ranked constraints requiring identity between input and output, such as the anti-deletion constraint MAXσ (McCarthy & Prince 1995) defined in (10).

(10) MAXσ Syllable Maximality: Every syllable in the input has a correspondent in the output.

The interaction between the prosodic constraints in (7) and the correspondence constraint in (10) is illustrated in the following tableau, where the input consists of five syllables.

(11) Word Minimality

Input: σσσσσ	FTBIN	PARSEσ	ALL-FT-L	ALL-FT-R	MAXσ
a. [(σσ)(σσ)σ]$_{PrWd}$		*!	**(2σ)	**(3σ,1σ)	
b. [σ(σσ)(σσ)]$_{PrWd}$		*!	****(1σ,3σ)	*(2σ)	
c. [(σσ)σ(σσ)]$_{PrWd}$		*!	***(3σ)	*(3σ)	
d. [(σσ)σσσ]$_{PrWd}$		*!**		*(3σ)	
e. [(σσσσσ)]$_{PrWd}$	*!				
f. [(σσσ)(σσ)]$_{PrWd}$	*!		***(3σ)	*(2σ)	
g. [(σσ)(σσ)]$_{PrWd}$			*!*(2σ)	*(2σ)	*
h. ☞ [(σσ)]$_{PrWd}$					***

[2] Consonants rendered unsyllabifiable due to final V-deletion are also deleted.

The input is located at the top left cell. Candidates are listed in the leftmost column in random order. The constraints are listed from left to right according to their ranking. Dotted lines indicate no ranking between the constraints and solid lines indicate crucial ranking. An asterisk indicates a violation of the specific constraint and an exclamation mark denotes a fatal violation of the constraint. Cells are shaded when a fatal violation occurs, or when only one candidate remains in the competition, and is thus the optimal candidate (indicated by ☞).

Ranking the prosodic constraints in (7) above MAXσ forces word minimality, regardless of the syllabic length of the input. For a candidate to remain faithful to an input comprising five syllables, as shown in tableau (11) above, it must violate at least one of the prosodic constraints. If all five syllables are parsed into a single foot, this would violate FTBIN (candidate (e)). If they are parsed into two binary feet, one syllable must be left unparsed (candidates (a)-(c)) or one foot must be ternary (candidate (f)). Either way, both alignment constraints are violated. To minimally violate the alignment constraints, by only violating one of them, more syllables must be left unparsed (candidate (d)). The only way to satisfy all of the prosodic constraints is by violating MAXσ (candidate (h)).

When undominated, ALL-FT-L/R will always select the candidate with a single foot. PARSEσ and FTBIN will ensure that this foot is binary without any syllables left unparsed. The output should optimally be the minimal word.

Note that while domination of faithfulness constraints by markedness ones forces word minimality, it also results in the minimal word being the maximal word. This property reflects some aspects of Semitic morphology, as well as a stage in language acquisition.

Smolensky (1996) argues that in the initial stage of acquisition the structural (markedness) constraints outrank constraints requiring faithfulness between input and output. Therefore, in this early stage, faithfulness to the input cannot incur violation of the structural constraints, and so early outputs are unmarked. The process of learning is characterised within OT as the reranking of the constraints, such that markedness constraints are demoted below faithfulness constraints, to allow more marked structures to surface.

However, as Piñeros (1999) points out, the faithfulness constraints which are dominated in the initial stage of acquisition require identity between input and output, while those dominated in abbreviations require identity between two output forms. Since Input-Output faithfulness (IO-Faith) and Output-Output faithfulness (OO-Faith) are independent constraint families (McCarthy & Prince 1995), they may be independently ranked with respect to the markedness constraints. Thus, abbreviation is not the reversal of the acquisition process, whereby faithfulness constraints are demoted in favour of markedness ones.

1.3. Language Background

1.3.1. Serbo-Croatian

The phonological structure most relevant to abbreviations is that of the syllable. As argued in Chapter 1, and as will be assumed throughout this study, the optimal shortened output is a binary foot. Feet are binary either moraically (in weight sensitive languages) or syllabically (in weight insensitive languages). Zec (1999) proposes that in addition to the standard binary foot, Serbo-Croatian also possesses tonal feet which play a significant role in the language's prosody.

I follow Zec's (1999) analysis of Serbo-Croatian feet. According to Zec, Serbo-Croatian is quantity sensitive with a trochaic foot system. Toneless feet consist of either two light syllables (12a) or a single heavy syllable (12b).

(12) Toneless Foot Inventory

a. $(\sigma_{\mu}\ \sigma_{\mu})_F$ b. $(\sigma_{\mu\,\mu})_F$

Uneven toneless feet, such as $(\sigma_{\mu\mu}\ \sigma_{\mu})_F$, are banned due to a TROCHAICQUANTITY (TROCHQU) constraint which requires that the two units comprising the trochaic foot be equal in size in terms of moraic count and tonal association.

Only vowels contribute to syllable weight, so that Cvv(C)[3] syllables are heavy while Cv(C) ones are light.

Tonal feet consist of either a heavy syllable linked to a high (H) tone (13a), or a light syllable linked to a high tone (13b).

(13) Tonal Feet

a. $(\underset{\text{H}}{\sigma_{\mu\,\mu}})_F$ b. $(\underset{\text{H}}{\sigma_{\mu}})_F$

TROCHQU also excludes disyllabic tonal feet such as $^{*}(\underset{\text{H}}{\sigma_{\mu}}\ \sigma_{\mu})_F$ from the Serbo-Croatian foot inventory.

Light syllables may be footed only when linked to a high tone (13b). The following data, drawn from Zec (1999), demonstrate the behaviour of the two classes of feet in the nominative case in terms of vowel length.

[3] There are no Cvv(C) AWs in Serbo-Croatian. If the base's initial segments form a Cvv(C) sequence, this sequence cannot form a proper foot, due to the high ranking of ONSET, and so an AW cannot emerge. In this case, the abbreviated form to emerge would most likely be a Clipped Compound (e.g. ekoloških organizaciya srbiye > *ékos* **éos* 'Ecology organisation of Serbia'). Also, a vowel cannot be lengthened for the AW to reach a binary foot status, because FAITHLENGTH outranks DEP-H (see (26)).

(14) a. Toneless Feet

Nom. sg.	Gen. sg.	Gloss.
brood	broda	'boat'
boog	boga	'god'
goost	gosta	'guest'
tvoor	tvora	'skunk'

b. Tonal Feet[4]

Nom. sg.	Gen. sg.	Gloss.
rák	ráka	'crab'
bég	béga	'bey'
brát	bráta	'brother'
grád	gráda	'hail'

The toneless bases in (14a) (represented by the base of the genitive forms) must undergo vowel lengthening in the nominative case to reach foot status (i.e. bimoraicity) and meet the minimality requirement whereby a prosodic word must minimally contain a foot.

The monomoraic bases in (14b), by contrast, are linked to a high tone, satisfying the foot prominence requirement, whereby a foot must be linked to a tone, and so no vowel lengthening occurs.

In Serbo-Croatian, therefore, in order to satisfy the minimal word requirement, a foot must satisfy either the Foot Binarity or the Foot Prominence condition.

1.3.2. Hebrew

Modern Hebrew shows no evidence of moraic structure, and is taken to be quantity insensitive. This assumption is based on the facts that stress assignment does not distinguish between open and closed syllables and there is no length contrast. Stress often appears on the final syllable, whether CV or CVC (e.g. miv.cá, miv.cár).

[4] The yer vowel pertaining to the nominative singular has been supressed, as it remains unfooted when at the right edge of the PrWd. X́ here denotes a high tone and not stress. In the Genitive case in both (14a) and (14b), the yer vowel is present, but remains unfooted if it is final in a prosodic word.

Consequently, all syllables are monomoraic and, therefore, prosodic words are minimally disyllabic. Existing monosyllabic words, such as *yad* 'hand', *kaf* 'spoon', and *dli* 'bucket' are remnants of the old system which was sensitive to weight (*kaaf*, *yaad*, *dəli*). New monosyllabic words are rarely introduced into the language (cf. *faks* 'fax'). Nevertheless, a few monosyllabic AWs involving a glide at the left edge of one of the non-initial base words do exist. I will claim that these are, in fact, the only pure AWs exhibited in the language.

1.3.2.1. Syllable Inventory. Modern Hebrew has a very limited inventory of syllable structures. Vowel length is not distinctive and tautosyllabic consonantal clusters are restricted. Complex codas are rather rare, while complex onsets, though much more common, are mostly found word-initially.

The range of possible syllable structures are: (C)V (*a.gaf* 'department'), VC (*ir.gun* 'organisation'), CVC (*mad.rix* 'instructor'), CCV(C) (*plu.ga* 'troop', *sgan* 'deputy') and in rare cases, also CVCC (*sport* 'sports') and CCCV(C) (*sprint* 'sprint').

1.3.2.2. Foot Inventory. Following Graff 2000, Hebrew has a trochaic foot system, with syllabic feet. In the nominal system, if no lexical pre-specifications intervene, main stress is assigned to the rightmost syllable:

(15) ga.(mád) 'dwarf'
tav.(lín) 'spice'
xa.(túl) 'cat'

Secondary stress alternates rhythmically and is assigned leftwards (Bolozky 1982) and degenerate feet are allowed at the right edge due to high priority of the final stress requirement (Graf 2000):

(16) me.(là.fe).(fón) ‘cucumber’
(mìt.ri).(yá) ‘umbrella’
(tà.a).(rí̋f) ‘rate’

1.3.3. English

1.3.3.1. Foot Inventory. English has a trochaic foot system, so that prominent (i.e. stressed) syllables are located at the left edge of the foot. Non prominent (i.e. stressless) open syllables undergo vowel reduction (stress is either primary or secondary):

(17) (mì.nə)(só.tə) Minnesota
(á.grɪ)(kʌ̀l.čə) agriculture
(gɛ̀.nə)(réɪ.šən) generation

The syllables of prosodic words are parsed into feet so that no more than one unparsed syllable may intervene between a foot and word boundary or between any two feet (Hammond 1997). This gives rise to the following parsing options for disyllabic and trisyllabic words:

(18) a. Disyllabic PrWds[5]
(σ́ σ) háppy
σ (σ́) abóut
(σ́)(σ́) rébòund
b. Trisyllabic PrWds
σ (σ́)(σ́) elέctrὸn
σ (σ́ σ) vaníl̋la
(σ́)(σ́)(σ́) chìmpànzée
(σ́)(σ́ σ) bàndána
(σ́ σ) σ Cánada
(σ́ σ)(σ́) Tènnessée

[5] (σ́)σ is excluded in Hammond’s analysis due to the binarity constraint which militates against monosyllabic feet followed by an unfooted syllable.

Trisyllabic words do not surface in the grammar of abbreviations due to the high ranking of the minimal word constraints (see section 1.2).

Though the disyllabic examples in (18a) have three attested parsings, only one, $[(\acute{\sigma}\ \sigma)]_{PrWd}$, surfaces in the grammar of abbreviations due to the high ranking of PARSEσ (which excludes $[\sigma\ (\acute{\sigma})]_{PrWd}$) and ALLFTR/L (which excludes $[(\acute{\sigma})\ (\acute{\sigma})]_{PrWd}$).

1.3.3.2. Syllable Inventory. English allows a wide range of possible syllables, from CV syllables (*má.nə.gə(r)* 'manager') through onsetless syllables *(ə.bóʊt* 'about') to syllables with complex onset, complex nucleus and complex coda (*stréɪnj* 'strange') all of which are attested in the abbreviations explored in the following chapters.

In the remainder of this study I show that the constraints responsible for the deletion of segments, i.e. the minimal word constraints, are the same for all four types of abbreviations in all three languages discussed. The thesis is organised as follows: In Chapter 2 I explore how Acronym Words (AWs) emerge from multi base words, and what constraints are responsible for the selection of the segments comprising the AW. In Chapter 3 I look at Clipped Compounds (CCs) and show that the same constraints which account for AWs can account for CCs as well. In Chapter 4 I show how bases comprised of a single word are shortened in much the same way as AWs and CCs. In Chapter 5 I look at Hypocoristics which are shortened personal names that more often than not involve a suffix. Finally, in Chapter 6 I compare the abbreviation types and point out the similarities as well as the differences between them, and between the three languages explored.

Chapter 2. Acronym Words

2.1. Introduction

Acronym Words (AWs) are single Prosodic Words (PrWds) formed from the initial segment of every PrWd in a multi-word source. In this chapter, I analyse the grammar of AWs in Serbo-Croatian (Section 2.2), Hebrew (Section 2.3) and in English (Section 2.4), and show how the formation of AWs is affected by the Minimal Word constraints discussed in the previous chapter.

2.2. Serbo-Croatian

The examples in (19), divided according to the number of words in the base, are representative:

(19)

	Base	AW	Gloss
a.	Savezno Izvršno Veće	sív	'Federative executive assembly'
	Sekretaryat Unutrašnih Poslova	súp	'Ministry of internal affairs'
	Yugoslovenski Aero Transport	yát	'Yugoslav air-transport'
	Narodno Oslobodilaćki Rat	nór	'War of national liberation'
	Apatinska Modna Obuća	ámo	'Apatin Footware Fashion'
	Elektro-Distribuciya Beograd	édi	'Belgrade electric distribution'
b.	Autonomna Kossovsko Metohijska Oblast	ákmo	'Autonomous Kossovo and Metohija Area'
	Akademsko Kulturno Umetnićko Društvo	ákud	'Academic cultural artistic company'
	Auto Motor Sport Klub	ámsk	'Automobile sports club'

c.	Antifašističko Veče Narodnog Oslobajenya Jugoslaviye	ávnoy	anti-fascistic council of national liberation of Yugoslavia
	Fabrika Armatura, Specijalnih Mašina i Alata.	fásma	'Factory for fossets, special machinery and tools'
	Penzionersko Amatersko Kolturno Umetnićko Društvo	pákud	'retirement artistic and cultural company'
	Beogradska Revija Amaterskih Malih Scena	bráms	'small amateur theatre organisation of Belgrade'

The output AW can either be monosyllabic or disyllabic, regardless of the number of words in the base. It cannot, however, exceed two syllables. It is formed from the first (and only the first) segment from each prosodic word in the base, so that no base word is ignored and the linear order of the segments as they appear in the input is preserved (i.e. reflecting the order of the words). The faithfulness constraints responsible for this are ANCHOR$L_{I\text{-}O}$ (20a), ANCHOR$L_{O\text{-}I}$ (20b), and LINEARITY (21):

(20) a. ANCHOR$L_{I\text{-}O}$ Every segment at the left edge of a PrWd in the input must have a correspondent in the output.

b. ANCHOR$L_{O\text{-}I}$ Every segment in the output must have a correspondent at the left edge of a PrWd in the input.

The input assumed here is the entire base comprised of prosodic words, rather than the segmental material extracted from the base to serve as input, as proposed by Bat-El (1994) for Hebrew AWs.

ANCHOR$L_{I\text{-}O}$, as defined in (20a), ensures that every source word has a representative in the output, and ANCHOR$L_{O\text{-}I}$ (20b) ensures that no other segment beyond the first of each word appears in the output. Combined, the two constraints

ensure that the output is an acronym word (if it survives the evaluation of the prosodic constraints FTBIN/FTSAL, PARSEσ, and ALL-FT-L), such that the number of segments in the output equals the number of prosodic words in the input and only the segment at the left edge of each prosodic word appear in the output.

(21) LINEARITY Every two segments in the output reflect the order of precedence of the corresponding segments in the input (and vice versa).

LINEARITY bans metathesis of elements standing in correspondence relations between input and output. In McCarthy and Prince (1994), LINEARITY is not limited to reflect the linear order of segments within a PrWd: "Any two elements of a string will stand in an order relation which is necessarily preserved under LINEARITY" (p.7). I apply here their definition in a way that "elements" = segments and "string" = input (in this case multi-word). Thus, any two segments in the output must reflect the order of precedence of the input. Note that LINEARITY does not imply adjacency.

The three constraints (20a), (20b), and (21) are never violated in the grammar of AWs and so are assumed to be undominated.

As discussed in Section 1.3.1 above, Serbo-Croatian has two types of feet: one adhering to the foot binarity constraint and the other to the foot prominence constraint. Both types satisfy the minimal word requirement. The constraint on prominence, FTSAL (22), requires feet to be associated with tone.

(22) FOOT SALIENCE (FTSAL) (Zec 1999)
A foot should be associated with tone.

The two constraints (FTBIN and FTSAL) are not crucially ranked and so a foot will always satisfy one and violate the other. Satisfaction of both constraints through

vowel lengthening is not possible due to FAITHLENGTH (25) (see tableau (26)).[6] The two constraints, in addition to being undominated, must therefore also remain crucially unranked with respect to one another in order to allow a larger variety of feet, both tonal and toneless.

Tableau (24) demonstrates the interaction of the constraints active in the grammar of Serbo-Croatian AWs. The relevant faithfulness constraint here is MAXSEG (23) which penalises for every segment loss, rather than every syllable, as in Chapter 1. The minimal word constraints, FTBIN, PARSEσ, and ALL-FT-L, have been combined for space reasons. In case of violation, the violated constraint is indicated.

(23) MAXSEG Segment Maximality: Every segment in the input has a correspondent in the output.[7]

(24)

savezno izvršno veće	MinWd	$\text{ANCHORL}_{\text{I-O}}$	$\text{ANCHORL}_{\text{O-I}}$	LINEARITY	MAX SEG
a. ☞ (sív)	*(FTBIN)				15
b. (sa.iv)	*(FTSAL)		*!(a)		14
c. (sa.izv)	*(FTSAL)		*!*(a,z)		13
d. (sa.viv)	*(FTSAL)		*!*(a,v)		13
e. (si.ve)	*(FTSAL)		*!(e)		14
f. (vís)	*(FTBIN)			*!	15
g. (vív)	*(FTBIN)	*!(s)	*(v)		15
h. (siv)	**!(FTBIN, FTSAL)				15

Candidates (b)-(e) and (g) are discarded as they all violate $\text{ANCHORL}_{\text{O-I}}$ (candidate (g) also violates $\text{ANCHORL}_{\text{I-O}}$). Candidate (f) violates LINEARITY and is excluded as well. The remaining candidates (a) and (h) both violate FTBIN, but they differ in tonal association. Candidate (a) is associated with a tone, satisfying FTSAL,

[6] Recall that in Serbo-Croatian only vowels contribute to weight.

[7] Violations of MAX are evaluated as "more" or "less", in the usual manner. Numbering is applied here due to lack of space and does not indicate counting in any way.

while candidate (h) violates this constraint by remaining toneless, leaving candidate (a), the sole survivor, the optimal output.

A candidate not considered in (24) but which satisfies FTBIN (by having a long vowel) as well as ANCHORL$_{I\text{-}O}$ and ANCHORL$_{O\text{-}I}$ and is therefore in competition with the optimal candidate (25a) is *si:v*. However, by having a long vowel in the output, this candidate violates FAITHLENGTH (25), which prohibits moras in the output that do not have correspondents in the input.

(25) FAITHLENGTH A long/short vowel in the input is long/short in the output.

Furthermore, the assignment of tone to the AW violates DEP-H, a constraint prohibiting high tones in the output that do not have correspondents in the input. FAITHLENGTH must therefore dominate DEP-H as it is better to insert a high tone in order to reach a permissible foot, than it is to lengthen the vowel (candidate (b)). FAITHLENGTH also outranks FTSAL and FTBIN, as it is better to violate one of them than to lengthen the vowel in order to satisfy both (candidate (c)). This is demonstrated in tableau (26)[8].

(26)

savezno izvršno veće	FAITHLENGTH	FTSAL	FTBIN	MAXSEG	DEP-H
a. ☞ (sív)			*	15	*
b. (si:v)	*!	*		15	
c. (sí:v)	*!			15	*

To generalise the discussion thus far, the hierarchy of the constraints involved in the formation of Serbo-Croatian AWs is such that the minimal word constraints, which in this language include FTSAL in addition to FTBIN, are ranked above faithfulness ones (except FAITHLENGTH, which would otherwise result in

[8] The non-violated constraints: PARSEσ, ALL-FT-L, ANCHORL, and LINEARITY, have been supressed for space reasons.

lengthening, which is ill-favoured in abbreviations), in accordance with the discussion in Chapter 1.

It is striking, however, that the prosodic words of the base of Serbo-Croatian AWs in (19) happen to have a correct distribution of consonants and vowels, so that possible syllables can be formed without having to add material, nor to ignore the initial segment of some base words. When the initial segments cannot form possible syllables in this way, more material can be extracted from the base (as in CCs to be discussed in Chapter 3). Hebrew, however, not only allows to extract more segmental material (section 2.3 and 3.3), but also provides some repair strategy to rescue these clusters and create possible syllables.

2.3. Hebrew

Hebrew Acronym Words can be divided into three groups.[9]

(27)	AW	Base	Gloss
a.	ʔábax	ʔatomi biyologi ximi atomic biological chemical	‘atomic biological & chemical (warfare)’
	cáhal	cva hagana le-israel army defence to-israel	‘Israeli Defence Force’
	natbág	nemal teufa ben gurion port flight Ben-Gurion	‘Ben-Gurion airport’
b.	ʔéšel	ʔoxel štiya (ve-)lina food drink (and) sleep	‘board and lodging’
	šékem	šerut kantinot (u)miznonim service canteen fem. pl. (and) buffets	‘canteen service’
	mécax	mištara cvaʔit xokeret police military fem. investigates fem.	‘military police investigation (unit)’
c.	pérax	proyekt xonxut project coaching	‘“big brother/sister” project’
	ʔerán	ʔezra rišona nafšit aid first sg. fem. mental sg. fem.	‘first aid for mental distress’

[9] Whether the glottals are phonetically realised or not has no bearing on the analysis.

ya?él yad ?ezer laxolim ‘helping hand for the sick’
hand help for the sick masc. pl.

Group (a) comprises the majority of Hebrew AWs of the shape Ca(C)CaC. Less common are group (b), referred to as “segolates”, which have the shape CéCeC (triconsonanatal disyllables with penultimate stress where the vowel is generally *e*, or *a* in the presence of historical gutturals). Group (c) where the AWs take the shape of existing words (*péra*x ‘flower’) or names (?erán, ya?él) is even less common.[10]

Bat-El (1994) assumes the written consonants of the Hebrew acronym to be the underlying input representation of the AW, as vowels are usually not marked in the Hebrew writing system.

(28)

Acronym	Acronym Base	Gloss
XK	Xaver Kneset	‘member of parliament’
MXT	Mefaked XaTiva	‘squadron commander brigadier’
PLMX	PLugot MaXac	‘shock troops’

Since the input is consonantal, vowels must be inserted for proper syllabification. The shape of the output is determined by the following set of constraints:

(29) Constraints

ONSET:	Syllables have onsets
NOCODA:	Syllables do not have codas
*COMPLEX:	No more than one C or V may associate to any syllable position node
PARSE:	Underlying segments must be parsed into syllable structure
MSEG:	Morphologically unsponsored segments are prohibited
ALIGNCODA:	Align R (C, PrWd)

Every inserted vowel incurs a violation of MSEG. The ranking of MSEG above NOCODA, ensures the minimisation in the number of syllables at the price of

[10] Not much can be said about type (c) except that if a consonantal string exists in the input as an existing string in the language, the vowels (as well as stress) are copied from the existing word to the AW. The considerations taken into acount here are not entirely phonological and therefore will not be dealt with in this paper.

violating NOCODA. *COMPLEX and ONSET restrict the possible syllables to CV and CVC.

These constraints are ranked as follows:

(30) Constraint hierarchy

PARSE, ONSET, *COMPLEX >> MSEG >> NOCODA >> ALIGNCODA

A demonstration of how this works is given in (31) (uppercase letters indicate the segments taken from the base to form the AW):[11]

(31) Acronym base: Mefaked XaTiva

/MXT/	PARSE	ONSET	*COMPLEX	MSEG	NOCODA	ALIGNCODA
a. ☞ ma.xat				**	*	
b. max.ta				**	*	*!
c. maxt			*!	*	**	
d. mxat			*!	*	*	
e. ma.xa.ta				***!		
f. am.xat		*!		**	**	*

Though they least violate MSEG, candidates (c) and (d) are ruled out for violating the undominated constraint *COMPLEX. Candidate (e) is ruled out because it incurs more violations of MSEG than the remaining candidates, even though it satisfies NOCODA. Candidates (a) and (b) equally violate NOCODA, but candidate (a) is finally selected as its coda is at the right edge satisfying ALIGNCODA while that of candidate (b) is not.

Under this analysis, all Hebrew AWs are instances of impossible strings of consonants and vowels (in this case *a*) must be inserted to allow correct parsing into syllables. In section 4.3 I propose a different analysis which assumes the entire acronym base to be the input and the constraint hierarchy to be responsible for the selection of segments forming the acronym word. This analysis reduces the

[11] See Bat-El (1994) for full discussions on the arguments for the constraint hierarchy in (30) and on the epenthetic vowel.

distinction between Hebrew and other languages, and within Hebrew, provides a unified account for AWs and CCs.

Hebrew is not unique in inserting vowels to save impossible clusters in the formation of AWs; English also has a very small group of such cases, as shown in the following section.

2.4. English

In Serbo-Croatian the initial segments of the PrWds comprising the base can form pronounceable words, whereas in Hebrew they seldom do. In this section, I look at English AWs, where the initial segments of some bases do form pronounceable words, while those of others do not. The latter are sometimes rescued using a similar vowel epenthesis strategy suggested for Hebrew AWs (section 2.3).

The following data illustrate the two groups discussed here:

(32) English Acronym Words

	Base	AW
a.	American Standard Code (for) Information Interchange	ASCII (æski:)
	North Atlantic Treaty Organisation	NATO (néytoʊ)
	Light Amplification (by) Stimulated Emission of Radiation	LASER (léyzər)
b.	West Coast Conference (on) Formal Linguistics	WCCFL (wíkfɪl)
	Formal Linguistic Society (of) Mid-america	FLSM (fʊ́lsʊm)

Here too, the two ANCHORL constraints (20a) and (20b) ensure that the first (and only the first) segment from each PrWd in the base comprises the output. These segments are parsed into syllables and feet to the satisfaction of the remaining constraints, as suggested in section 2.2, and illustrated in the following tableau:

(33)

north atlantic treaty organisation	FTBIN	PARSEσ	ALL-FT-L	ANCHORL I-O	ANCHORL O-I	LIN.	MAXSEG
a. ☞ (na.to)							27
b. (nat.ro)					*!(r)		26
c. (no.to)				*!(a)	*(o)		27
d. (no.ra)to		*!			*!*(o,r)		25
e. (nat)				*!(o)			28

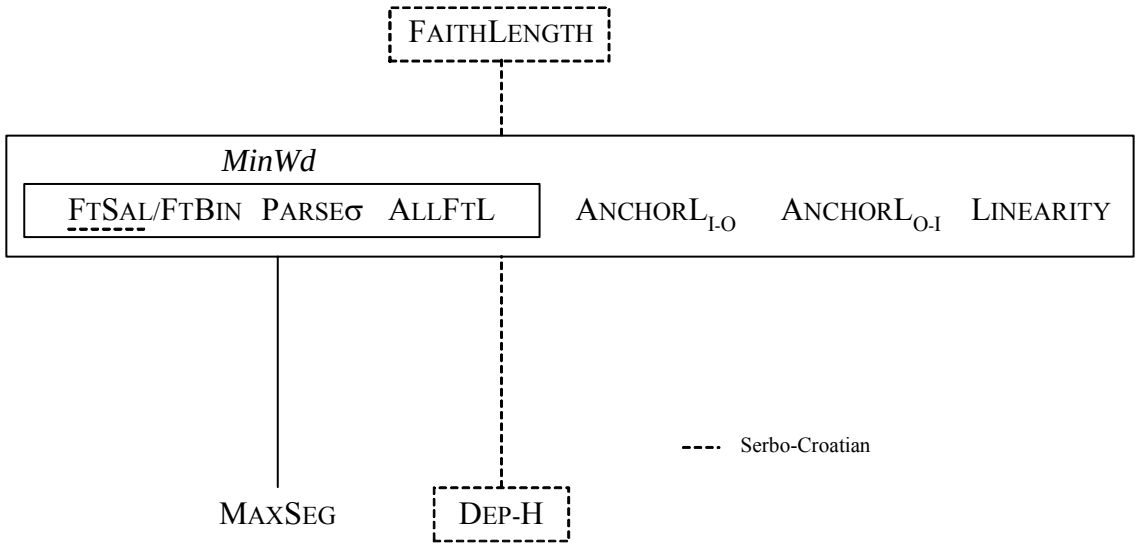
All candidates except for candidate (a) violate at least one of the undominated constraints, leaving (a) as the winning candidate.

Note that for the (32b) cases, vowels are inserted to rescue the unsyllabifyable clusters, even though this violates MSEG, which prohibits the presence of unsponsored morphological segments. However, this rescue strategy is quite rare in English.

2.5. Conclusion

The grammar of abbreviation into AWs for Serbo-Croatian and for English involves the constraint hierarchy in (34), where the minimal word constraints, along with ANCHORL and LINEARITY outrank the faithfulness constraints, except FAITHLENGTH. FAITHLENGTH needs to be ranked above the minimal word constraints to avoid lengthening, which is ill-favoured in the grammar of abbreviation. In fact, lengthening is not attested in abbreviations in any of the languages discussed here.

(34) Constraints Hierarchy



The undominated faithfulness constraints are ANCHORL$_{I-O}$, ANCHORL$_{O-I}$, which are responsible for the selection of the segments at the left edge, and LINEARITY, which militates against reordering of segments.

FTSAL, and DEP-H, the discussion of which was raised for Serbo-Croatian (indicated by a dotted line in (34) above) are also ranked such that the prosodic constraints (FTSAL) outrank the faithfulness constraint (DEP-H). FAITHLENGTH is not violated in any of the abbreviation types discussed hereinafter. Its presence so high in the hierarchy is irrelevant to their analysis, and so it will no longer be mentioned.

The constraints involved in Hebrew AWs are left out of the scheme in (34). I return to them in Chapter 3.

Chapter 3. Clipped Compounds

A Clipped Compound (CC) is a single pronounceable word formed from the beginnings of each of the words comprising the multi-word base. AWs and CCs are similar in this respect as they both have multi-word bases and both are minimal words. But CCs emerge when AWs cannot, because the initial segments from each base word are not enough to form a minimal word.

In this section I will show that in the grammar of CCs too, the minimal word determines the number of segments needed from the input, which is "just enough for a minimal word".

3.1. Serbo-Croatian

Recall that in section 2.2 on Serbo-Croatian AWs, the initial segments of the base words showed a correct distribution of consonants and vowels. In (35) below, however, the initial segment of each source word is not enough to form an acceptable PrWd as in section 2.2, since these segments cannot be properly footed.

The following examples are representative.

(35) Serbo-Croatian CCs

Base	Clipped Form		Gloss
EKoloških Organizaciya Srbiye	EKOS	*eos	'Ecology organisation of Serbia'
Industriya Grajevinskog MAteriyala	IGMA	*igm	'Industry of building material'
KOMunalno GRAjevinsko Preduzeće	KOMGRAP	*kgp	'Community building company'
GRAjevinski METal	GRAMET	*gm	'Metal industry'
FAbrika MOtora Sarayevo	FAMOS	*fms	'Sarajevo motor plant'

The initial segments of the base words, represented by the asterisked forms, do not form acceptable feet: **kgp, *gm,* and **fms* do not have proper nuclei; **igm* violates the SONORITY constraint (Hammond 1997), whereby within the syllable, onsets must increase in sonority and codas must decrease in sonority; and **eos* violates the high ranked ONSET constraint.

As suggested throughout this study, truncation is possible when the prosodic constraints responsible for word minimality are ranked above constraints which require identity between input and output. This is demonstrated for clipped compounds in tableau (36) below.

Recall from Chapter 2 that there are two ANCHORL constraints: between input and output and between output and input. The former requires every segment at the left edge of every PrWd in the input to have a correspondent in the output, and the latter, every segment in the output to have a correspondent at the left edge of a PrWd in the input. Neither constraint is violated in AWs, but in CCs, because more segments from each base word are needed, $\text{ANCHORL}_{\text{O-I}}$ is often violated, while $\text{ANCHORL}_{\text{I-O}}$ is not. Therefore, $\text{ANCHORL}_{\text{I-O}} >> \text{ANCHORL}_{\text{O-I}}$.

$\text{ANCHORL}_{\text{I-O}}$ is not violated and is therefore assumed to be undominated. $\text{ANCHORL}_{\text{O-I}}$ is dominated as it can be violated in order to satisfy FTBIN. $\text{ANCHORL}_{\text{O-I}}$ also outranks MAXSEG. These two constraints are in competition with one another as $\text{ANCHORL}_{\text{O-I}}$ requires the output to be as short as possible (every segment other than the initial of every base word incurs a violation of the constraint) while MAXSEG requires it to be as long as possible (by penalising for every deleted segment).

Here, too, LINEARITY ensures that the order of precedence in the ouput is as in the input. It should be noted, however, that here, LINEARITY refers to the precedence relations within the PrWd, as well as between PrWds, as seen in the analysis of AWs.

As argued in Chapter 2, adjacent segments in the input are not necessarily adjacent in the output and vice versa.

(36)

fabrika motora sarayevo	MinWd	CONTIG.	LIN.	ANCHORL I-O	ANCHORL O-I	MAXSEG	ALIGNC ODA
a. (fms)	*!(FTBIN)					18	
b. (fams)	*!(FTBIN)				*	17	
c. (fa.mo)				*!	**	17	
d. ☞ (fa.mos)					**	16	
e. (fab.mos)					***!	15	*
f. (fam.sa)					**	16	*!
g. (fa.mo)sa	*!(PARSE)				***	15	
h. (fa.sam)			*!		**	16	
i. (fi.mos)		*!			**	16	

Candidates (a) and (b) violate FTBIN as they are not bimoraic and are therefore ruled out (recall that in Serbo-Croatian only vowels are moraic). Candidate (c) is ruled out by $\text{ANCHORL}_{\text{I-O}}$ as one of the base words is not represented in the output. The three candidates (d-f) all satisfy the undominated constraints, but candidate (e) has more violations of $\text{ANCHORL}_{\text{O-I}}$ than the other two candidates (d) and (f), which equally survive the competition. The optimal candidate (d) is selected by a lower ranked constraint, ALIGNCODA (29) as in the Hebrew AWs in Chapter 2 (see (31)).[12]

CONTIGUITY (37) is required to rule out candidates comprising segments which are not contiguous in the input (candidate (i) in (36) above).

(37) CONTIGUITY: A contiguous string in the input must correspond to a contiguous string in the output.

CONTIGUITY, specifically INPUT CONTIGUITY (I-CONTIG) (McCarthy and Prince 1995) rules out deletion of elements internal to the input string. Thus, xyz → xz

[12] A candidate not considered here is the monosyllabic *fáms* (associated to a tone). Given the constraint hierarchy in section 2.2, this candidate should be selected by $\text{ANCHORL}_{\text{O-I}}$ as the winner. This, however, is not the case in any of the Serbo-Croatian CCs, which suggests that DEP-H is undominated, and will, therefore, not be considered further.

violates I-CONTIG, as xz is not a contiguous string in the input. However, deletion at the edges is ok: xyz → xy, as xy is a contiguous string in the input.

It should be noted here that although CONTIGUITY has not been applied in the analysis of AWs, its presence so high in the hierarchy of constraints does not pose a problem to their analysis. This constraint seems to refer to the base PrWd, and so this constraint can never be violated in AWs. In the analysis of CCs, it rules out candidates that select non-contiguous segments in the input.[13] Thus, *komgrap* (KOMunalno GRAjevinsko Preduzeće) is better than *komgap since *gap* is not a contiguous string in the input, even though the latter contains a complex onset which may be disfavoured in the grammar of abbreviations as it is more marked than a non-complex onset.

3.2. English

English CCs formed from two word bases, consist of two heavy syllables, as demonstrated in (38) below.

(38) English CCs

Base Words	Clipped Form	
SITuation COMedy	sitcom	*sic
AViation GASoline	avgas	*ag
FORmula TRANslation	fortran	*fot
PARallax SECond	parsec	*pas
REDuction OXidation	redox	*redo
MASs CONcentration	mascon	*mac
NATional COMmunications	natcom	*nac
WINdows MAGazine	winmag	*wim
AMphibious TRACtor	amtrac	*at

[13] It is, however, violated in Heb CCs (see Section 3.3).

The constraints responsible for the shape of all abbreviations discussed thus far, essentially select the shortest allowed foot. In the analysis above, the constraint responsible for the minimal size of the foot is ANCHORL$_{O\text{-}I}$ (20b) ranked above MAXSEG (23). However, the English CCs in (38) are not the shortest acceptable feet which can be generated, suggesting a different ranking, such that ANCHORL$_{O\text{-}I}$ is outranked by MAXSEG (tableau (39)).

(39) Constraints interaction in English CCs

situation comedy	MINWD	CONTIG.	LIN.	ANCHORL$_{I\text{-}O}$	MAXSEG	ANCHORL$_{O\text{-}I}$
a. (sic)					12!	*
b. (si.co)					11!	**
c. (sit).co	*!(PARSE)				10	***
d. (sit).(co)	*!*(FTBIN,ALLFTL)				10	***
e. (sit.co)					10!	***
f. (sit).(com)	*! (ALLFTL)				9	****
g. ☞ (sit.com)					9	****

Candidates (c), (d), and (f) violate one or more of the MinWd constraints and are therefore disqualified, leaving candidates (a), (b), (e), and (g). All things being equal, MAXSEG selects the most faithful of the group, candidate (g).

The grammar of English CCs is essentially that of AWs, discussed in Chapter 2 where the initial segments from each base word in most cases are not enough to form a proper foot. For English CCs, ANCHORL$_{O\text{-}I}$ is ranked below MAXSEG, setting it apart from both AWs and the Serbo-Croatian CCs in 3.1.

3.3. Hebrew

Returning to the analysis of Hebrew AW, recall that Bat-El's (1994) analysis assumes a consonantal input (from written material). In this section, I propose that the same constraints presented in section 3.1 (for Serbo-Croatian) and 3.2 (for English), are responsible for the selection of the segments involved in CCs, as well as for the prosodic shape of the surface CCs, and that in essence, the Hebrew data in

(27) and (28) are not AWs at all, but CCs. In the Hebrew writing system, AWs (as well as regular acronyms) and CCs are graphically marked by double quotes between the two rightmost consonants. Hebrew speakers and grammarians regard both AWs and the CCs discussed hereinafter as AWs and do not differentiate the two, suggesting that for them the two types are one and the same. I argue that they are indeed the same, but also that they are CCs rather than AWs.

First, we turn to CCs with bases comprising three PrWds. Note that neither bases nor output forms differ from the AWs discussed in section 2.3, the differences lie in the proposed analyses.

(40) 3-Words Bases

Base	CC	Gloss
ʔatomi biyologi ximi	ʔábax	'atomic biological chemical (warfare)'
ʔanašim xašuvim meod	ʔaxám	'very important people (VIP)'
xavurat pikud kidmit	xapák	'front command unit'
pikadon kcar moed	pakám	'short term deposit'

ANCHOR$L_{I\text{-}O}$ ensures that the initial segment from each base word is represented in the output form. Because the initial segments of the base words in Hebrew are mostly consonants, more segments are required for proper syllabification. The ANCHOR$L_{I\text{-}O}$ constraint (20a) allows the extraction of more than the initial segments, as long as the initial ones are extracted as well.

Observe that the input suggested here differs from Bat-El's consonantal input and consists of full structured PrWds comprising the acronym base, as in English and Serbo-Croatian. The prosodic constraints, FTBIN, PARSEσ and ALL-FT-L motivate the truncation, and the other constraints cited here are responsible for the selection

and distribution of the segments forming the CC, much like in the other two languages discussed in this chapter.

But Hebrew differs from English and Serbo-Croatian in the quality of the CC's vowels. Both English and Serbo-Croatian preserve the base vowels in CCs as well as in AWs. In Hebrew, the vowel *a* (the unmarked vowel) takes over in most cases.

The constraint responsible for this is stated in (41).

(41) $*V_{[\neq low]}$ No vowels other than *a* are allowed.

The constraint in (41) is drawn from a family of constraints, banning all vowels ranked according to the sonority scale, i.e. $*V_{[high]} >> *V_{[mid]} >> *V_{[low]}$. For brevity I combine the first two constraints: $*V_{[\neq low]} >> *V_{[low]}$. This set of constraints is ranked above IDENTF, which requires identity in feature values between corresponding segments.

The following tableau illustrates the type of CCs discussed in this section. LINEARITY and CONTIGUITY have been suppressed and candidates violating them, which never win, have not been considered.

(42)

xavurat pikud kidmit	MINWD	ANCHORL I-O	ANCHORL O-I	*V[≠low]	MAXSEG	IDENTF
a. (xa.pa)		*!	**		14	*
a. (xa.pik)			**	*!	13	
b. ☞ (xa.pak)			**		13	*
c. (xav.pak)			***!		12	*
d. (xa.pa)ka	*!(PARSEσ)		***		12	**
e. (xa.pa)(kad)	*!(ALL-FT-L)		****		11	**

Next, I turn to bases comprised of two prosodic words (43).

(43) 2-Words Bases[14]

Base	CC	Gloss
karov labait	kaláb	close to home
yexidat binuy	yaxáb	construction unit
mate klali	matkál	general headquarters
mefaked basis	mabás	base commander

The CCs in (43) are the shortest permissible foot (which recall, in Hebrew is disyllabic). The initial segment from each of the base words is not enough to form a proper foot, due to FTBIN. In addition to the proposed constraints, Hebrew has a FINALC constraint (44) (McCarthy 1993) which requires a consonant at the right edge of the prosodic word.

(44) FINALC PrWds end in a consonant.

(45)

yexidat binuy	MINWD	ANCHORL I-O	FINALC	ANCHORL$_{O\text{-}I}$	*V[≠low]	MAXSEG
a. (yab)	*! (FTBIN)			*		9
b. (ya.ba)			*!	**		8
c. ☞ (ya.xab)				***		7
d. (yax.ban)				****!		6
e. (yax.ba)			*!	***		7

FINALC rules out candidates with the shape Cv.Cv (45b), which otherwise least violate the higher ranked constraints. It also discriminates between Cv.CvC (45c) and CvC.Cv (45e) candidates which otherwise equally violate the undominated constraints.[15]

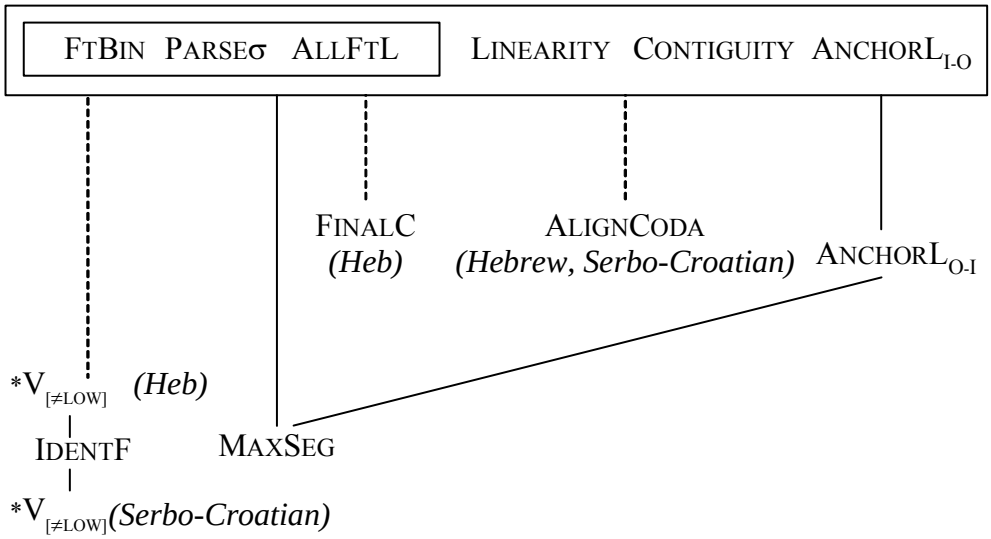
To summarise thus far, the constraints proposed here are ranked as follows:

[14] Note that in *matkal* < mate klali, CONTIGUITY is violated to avoid a cluster in the coda. This, however, does not explain why the surface form is not **matak* where CONTIGUITY is satisifed and the cluster is avoided without vowel epenthesis. There are several such cases: *mankal* < menahel klali 'general manager', *mafkal* <mefakeax klali 'inspector general', *mazkal* <mazkir klali 'secretary general'.

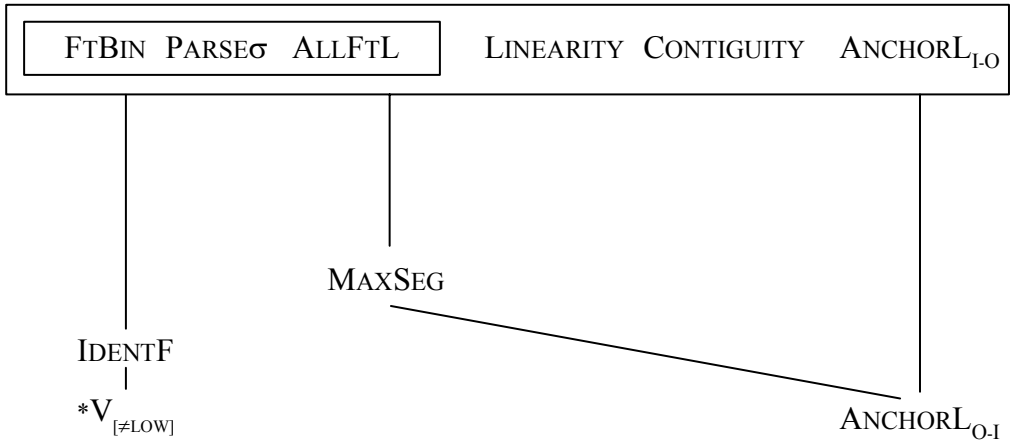
[15] Note, however, that the suggested analysis does not distinguish *yaxa͡b* from **yaba͡n*. Other such cases exist: *ama͡n* vs. **aga͡m* < agaf modiin 'Dept. of intelligence'; *rala͡š* vs. *raša͡l* < roš liška 'Head of office'.

(46) Constraints ranking

For Serbo-Croatian and Hebrew:



For English:



The ranking *V$_{[\neq low]}$ >> IDENTF is specific for Hebrew to account for the presence of *a* instead of the vowel inherent in the base. For English and Serbo-Croatian, the reverse ranking is observed: IDENTF >> *V$_{[\neq low]}$, so that the base vowel emerges in the output. Similarly, the presence of FINALC so high in the hierarchy is also particular to Hebrew CCs, though English CCs all end in a closed syllable (i.e. they end in a consonant) too. However, that AWs (e.g. *ascii, nato*) do end in vowels suggests that FINALC is ranked lower in English, as it is in Serbo-Croatian. English

differs from the other two languages in ranking ANCHOR$L_{O\text{-}I}$ below MAXSEG. This is summarised in the following table.

(47)	English	Hebrew	Serbo-Croatian
Common Ranking	FAITHLENGTH >> MinWd >> MAXSEG		
Conflicting Ranking	MAXSEG>>ANCHOR$L_{O\text{-}I}$	ANCHOR$L_{O\text{-}I}$ >> MAXSEG	
	IDENTF >> *$V_{[\neq LOW]}$	*$V_{[\neq LOW]}$ >> IDENTF	IDENTF >> *$V_{[\neq LOW]}$
	DEP-H >> FTSAL		FTSAL >> DEP-H
	ANCHOR$L_{O\text{-}I}$>> FINALC	FINALC>> ANCHOR$L_{O\text{-}I}$	ANCHOR$L_{O\text{-}I}$>> FINALC

Next, I discuss clippings (Chapter 4) and hypocoristics (Chapter 5), where in Hebrew, the base vowel emerges in the output abbreviation, suggesting the ranking IDENTF >> *$V_{[\neq LOW]}$ for Hebrew as well.

Chapter 4. Clippings

In Chapter 2 and Chapter 3 I have shown the conditions under which AWs and CCs are created from multi-word bases. Specifically, there are prosodic constraints that require that the output have a certain shape. When these constraints are ranked above constraints requiring identity between input and output, truncation may occur.

Within OT, restrictions are on the output, not the input, so that the input can contain any number of words. This chapter deals with clipping, a process whereby a single word is shortened forming a new prosodic word. I will look at English clippings[16] and show that they result from roughly the same constraints as those active in AWs and in CCs, when the input is a single word.

Bauer (1983) points out that the way in which the base word is shortened is unpredictable, because it is impossible to predict the number of syllables which will be retained, whether the final syllable will be open or closed or whether the stressed syllable will be included in the shortened form or not. The examples in (48) illustrate the problem pointed out by Bauer:

(48) English Clippings

	Base Word	Clipped Form	Base Syllables	
a.	memorándum	mémo	4	2 syllables retained, final syllable open, stressed syllable in the base is deleted.
	hippopótamus	híppo	5	
	metropólitan	métro	5	
	condomínium	cóndo	4	
b.	fratérnity	frát	4	1 syllable retained, syllable is closed, stressed syllable is deleted.
	gymnaèsium	gým	3	
	àdvertísement	ád	4	
c.	díscotheque	dísco	3	2 syllables retained, final syllable is open, primary stressed syllable is retained.
	phótograph	phóto	3	
d.	exàmination	exám	5	

[16] Hebrew has few clippings that look like hypocoristics in that they end in a suffix, and SC has none at all.

The data in (48) indicate that sometimes the syllable containing primary stress is preserved (*phóto* < *phótograph; dísco* < *díscotheque*), and sometimes not (*mémo* < *memorándum, frát* < *fratérnity*); and that regardless of the syllabic length of the base (whether 3, 4, or 5 syllables), the clipped form is maximally disyllabic, and if monosyllabic, as in (b), it must be minimally bimoraic, i.e. the unmarked minimal word, so that **fra*, **gy*, **a* are unacceptable abbreviations of *fraternity*, *gymnaesium*, and *advertisement* respectively.

The same minimal word constraints (FTBIN, PARSEσ, and ALL- FT-L), which are responsible for the abbreviation of AWs and CCs discussed in the previous chapters are at play here as well. However, the ranking of the constraints suggested in (46) accounts for only part of the data, as shown in tableaux (49) and (50).

(49)

advertisement	FTBIN ALL- FT-L	ANCHORL$_{I\text{-}O}$	ANCHORL$_{O\text{-}I}$	MAX SEG
a. a	*!(PARSEσ)			12
b. ☞ (ad)			*	11
c. (ad).ve	*!(PARSEσ)		***	9
d. (ad).(ve)	*!*(FTBIN, ALLFTL)		***	9
e. (ad).(ver)	*! (ALLFTL)		****	8
f. (ad).(vert)	*!(ALLFTL)		*****	7
g. (ment)		*!	****	9
h. (tise).(ment)	*!(ALLFTL)	*	********	4
i. (ad.vert)			**!***	7

Candidates (d), (e), and (f), though more faithful than the winning candidate (b), are discarded for having more than one foot. Candidate (d) also violates FTBIN as one of its feet contains less than two moras. Candidates (a) and (c) both contain an unparsed syllable, thus fatally violating PARSEσ. Candidates (g) and (h) attempt to preserve the ending of the base, but are discarded for violating ANCHORL$_{I\text{-}O}$. Finally, though there are candidates which are more faithful to the base form, candidate (b) is

selected as the optimal candidate as it does not violate any of the undominated constraints.

Ranking ANCHORL$_{O\text{-}I}$ above MAXSEG ensures that the shortest of the candidates which survive the undominated constraints ((b) and (i)) is selected as the optimal output. However, the following tableau reveals that this is not always the case:

(50)

memorandum		MINWD	ANCHORL$_{I\text{-}O}$	ANCHORL$_{O\text{-}I}$	MAXSEG
a.	me	*!		*	8
b.☞*	(mem)			**	7
c.+	(me.mo)			***!	6
d.	me.(mor)	*!*		****	5
e.	(me).(mor)	*!*		****	5
f.	(me.mo).ra	*!		*****	4
g.	(me.mo.ra)	*!		*****	4
h.	(me.mor)			***!*	5

Candidates (a) and (d-g) are eliminated after fatally violating one or more of the three higher ranked constraints: FTBIN, PARSEσ, ALL-FT-L. Candidates (b) and (c) both survive this evaluation. These are the bimoraic (b) and disyllabic ((c) and (h)) candidates. The ranking of ANCHORL$_{O\text{-}I}$ above MAXSEG, ensures the selection of the shortest acceptable foot (as in (49)), and so candidate (b) is wrongly selected as the optimal output (indicated by '☞*') instead of candidate (c), the actual output (indicated by '+'). I propose, therefore, that there are two main groups of clippings in English:

(51) Group 1

Base	Clip	
fraternity	frat	*frater
gymnasium	gym	*gymnas
advertisement	ad	*advert[17]

[17] In British English *advert* is the normative abbreviation, indicating that it belongs to Group 2.

sister	sis	*sist
trigonometry	trig	*trigo

(52) Group 2

Base	Clip	
hippopotamus	hippo	*hip
discotheque	disco	*dis
examination	exam	*ex
delicatessen	deli	*del
logotype	logo	*log

Group 1 seems to prefer the shortest allowed foot (i.e. the shortest minimal word), which is typically monosyllabic, while group 2 seems to prefer a longer foot, typically disyllabic, more faithful to the input base.

Nevertheless, though Group 2 prefers longer feet, the two groups do not differ in the ranking of ANCHORL$_{O\text{-}I}$ above or below MAXSEG, as seen in the discussion on English CCs (section 3.2). Rather, the difference between the two groups stems from the difference in the type of foot required: a moraic foot (FTBIN$_{\mu}$) for group 1 (as in tableau (49)) and a syllabic foot (FTBIN$_{\sigma}$) for group 2 (as in (53)).

(53)

memorandum	FTBIN$_{\sigma}$	PARSEσ	ALL- FT-L	ANCHORL$_{I\text{-}O}$	ANCHORL$_{O\text{-}I}$	MAXSEG
b. (mem)	*!				**	7
c. ☞ (me.mo)					***	6
h. (me.mor)					****!	5

Of the three candidates that survive the undominated constraints in (50), the monosyllabic one is discarded for violating FTBIN$_{\sigma}$, and the shortest of the two di-syllabic candidates (candidate (c)) is selected by ANCHORL$_{O\text{-}I}$, which outranks MAXSEG. Thus, while clips look like CCs in that they are formed from more than a single segment from the base, in effect, they exhibit the same grammar of AWs, where ANCHORL$_{O\text{-}I}$ outranks MAXSEG. Nevertheless, some clips (group 2) differ from AWs in that they require the output to be di-syllabic.

ANCHOR$L_{I\text{-}O}$ ensures that it is the left edge of the base word that is preserved.[18] A possible candidate for a tri-syllabic input $\sigma_1\sigma_2\sigma_3$ could, however, be a parsed foot comprised of the first and third syllables deleting material from the middle rather than from the right edge ($\sigma_1\sigma_3$). CONTIGUITY, defined in (37), prevents the deletion of any material from the middle, and LINEARITY ((21), repeated for convenience in (54) below), maintains the order of precedence as in the input (McCarthy and Prince 1995).

(54) LINEARITY: Every two segments in the output reflect the order of precedence of the corresponding segments in the input (and vice versa).

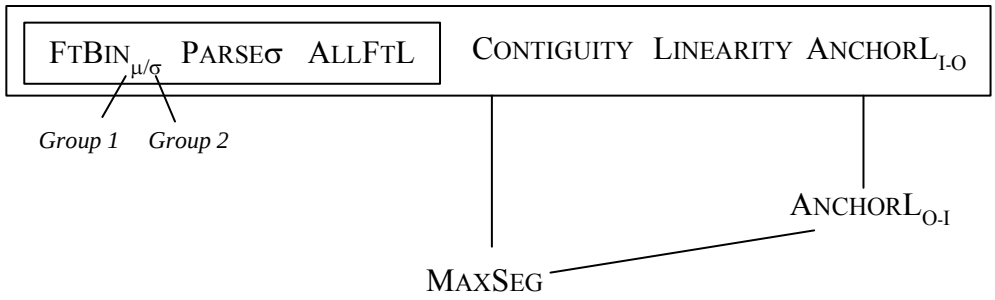
These three constraints ensure that the material retained from the base is taken from the beginning (i.e. the ends are dropped) (ANCHOR$L_{I\text{-}O}$), that no material is skipped (CONTIGUITY) and that the position of the segments relative to each other is also maintained (LINEARITY). Neither CONTIGUITY nor LINEARITY are violated in the grammar of clippings and so are assumed to be undominated.

In this chapter we saw that essentially the same grammar suggested in Chapter 2 for AWs and Chapter 3 for CCs holds for clippings as well.

CONTIGUITY needed to be introduced for clippings (as well as for CCs), which is absent in the analysis of AWs. Indeed, this constraint cannot be active in the grammar of AWs, because the segments that form the AW are never contiguous strings in the input. Nevertheless, this constraint is never violated in neither AWs nor CCs. The following diagram summarises the constraint hierarchy for English clippings.

[18] Some exceptions do exist: (in)flu(enza), (di)still(ery), and (re)frig(erator), where material is deleted from both edges, in violation of ANCHOR$L_{I\text{-}O}$.

(55) Constraints Hierarchy



In the following chapter, I turn to hypocoristics.

Chapter 5. Hypocoristics

Hypocoristics have drawn much attention compared to any other truncation form in a variety of languages (Mester 1990, Poser 1990 for Japanese; Itô and Mester 1997 for German; Davis and Zawaydeh 1999 for Arabic; Piñeros 1999 for Spanish, Weeda 1992, Benua 1995).[19] These studies all assume the source form for hypocoristic formation to be the full (surface) name. In correspondence terms, the claim is that in the formation of hypocoristics, correspondence is between two output forms rather than between input and output. This has been assumed in previous chapters to be the case for the other truncation processes as well.

The arguments supporting this claim are rather compelling. In some English hypocoristics, the vowel seems to be faithful to the output name even though this may violate some general constraints in the language that seem to hold elsewhere (Benua 1995; see also section 5.3.1 below):

(56) English Hypocoristics

[hæ.ri] > [hær]	*[hɑr]	Harry
[læ.ri] > [lær]	*[lɑr]	Larry
[sæ.ra] > [sær]	*[sɑr]	Sarah

Furthermore, in Spanish, one type of hypocoristic form remains faithful to the stressed syllable, though stress is not specified underlyingly (Piñeros 1999):

(57) Spanish Hypocoristics

seβastjáan > čáno

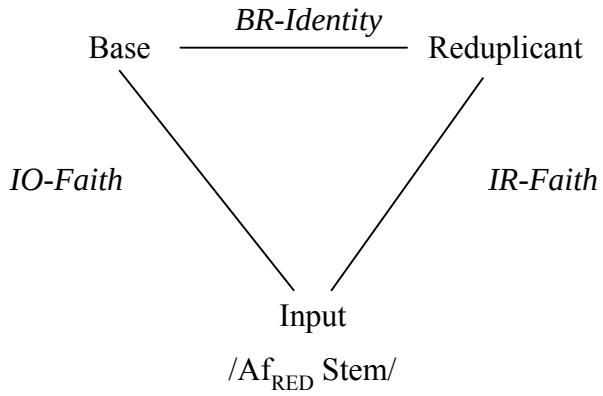
armándo > mándo

felísito > líčo

[19] Traditionally, hypocoristics and clippings have been treated as the same phenomenon. However, since clippings have been shown to be AWs with one base word and since hypocoristics are predominantly associated with a suffix, the two have been separated to facilitate the discussion.

In their study of reduplication, McCarthy and Prince (1995) propose a model of reduplication that calls for three correspondence relations: input-output faithfulness (IO-Faith), base-reduplicant identity (BR-Identity), and input-reduplicant faithfulness (IR-Faith).

(58) Reduplication model (McCarthy and Prince 1995)



The input, according to this model, is a stem and an empty RED affix, into which material is copied from the base subject to BR-Identity. Correspondence between input and reduplicant (IR-Faith) is necessary in cases where the reduplicant is more faithful to the input than the output base (see McCarthy and Prince 1995 for a detailed discussion), as in (59).

(59) Klamath Distributive Reduplication

a. Syncope in Base

/DIST + mbody’ + dk/	mbo-mpditk	‘wrinkled up (dist.)’
/DIST + smo’oq’y + dk/	sm’o-smq’itk	‘having a mouthful (dist.)’
/DIST + pniw + abc’ + a/	pni-pno:pc’a	‘blow out (dist.)’

b. Reduction in Base

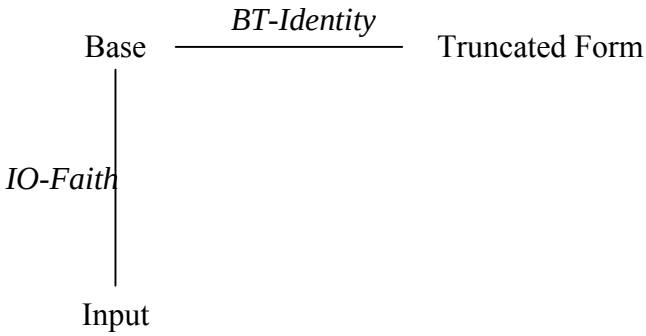
/DIST + sipc + a/	si-səpca	‘put out a fire (dist.)’
/DIST + Gatdk’ +a/	Ga-Gəttk’a	‘are cold (dist.)’
/DIST + pikca + ‘a:k’/	pi-pəkcaʔa:k	‘little pictures (dist.)’

In Klamath, the first vowel is deleted in closed syllables (a) or reduced in open syllables (b), if it is preceded by at least one syllable (Clements and Keyser 1983).

But while the vowel is deleted from the base in (a) or reduced in (b), the reduplicant prefix contains the input vowel, resulting in non-identity between the base and the reduplicant. In these cases, there must be correspondence between the reduplicant and the input.

Benua (1995, 1997) proposes to extend this model beyond base-reduplicant and to apply it to truncations as well:

(60) Truncation model (Benua 1995, 1997)



This model assumes only two correspondence relations instead of three, so that there is no correspondence relation between input and the truncated form. This predicts that the truncated form will never be more faithful to the underlying form than the base is. In section 5.3.2 I claim, however, that some English hypocoristics are faithful to the underlying vowel rather than to the base's surface vowel. Therefore, Benua's model will need to be revised to allow correspondence between the input UR and the truncated from (IT-Faith).

Generally speaking, hypocoristics are sometimes associated with a suffix ([-i] in Hebrew: odéd > ódi, binyamín > bíni, efráim, éfi; [-i] in English: jénɪfər > jéni, rɔ́bərt > rɔ́bi, wílɪəm > wíli; [-a], [-ka], [-ko], or [-an] in Serbo-Croatian: yélena > yéla/yélka, míroslav > mírko/míran), but not always: šošána > šoš (Hebrew), sámuel > sám (English), etc.). In the following analysis of hypocoristics, I adopt the OT

version of item-and-process approach, whereby affixes are viewed as constraints. Advocates of this approach include Russell (1995, 1999), Yip (1998), Hammond (1995), Bat-El (1999, 2000b), Adam and Bat-El (2000), Adam (to appear) (see also Hockett 1954, Aronoff 1976, Anderson 1992, for pre-OT works, in which affixes are viewed as rules).

Following Adam and Bat-El (2000) and Adam (to appear), the input for affixation is the stem specified for the required morphological category (e.g. nafal-[past 1st sg.]). The affix is introduced by an alignment constraint which places the affix at the specified edge of the stem and matches its morphological category to that required by the stem:

A&M[SUFF] Align and Match (stem-$_{[\text{PAST 1ST SG.}]}$, R, [ti]$_{[\text{PAST 1ST SG.}]}$, L)
Align the right edge of the stem with the left edge of the suffix and match the category of the suffix with the one required by the stem.

When the morphological constraint outranks faithfulness constraints (e.g. ANCHORRσ which requires the right edge of the input to correspond to that of the output), the result is *nafalti*, as shown in tableau (61):

(61)

Input: nafal-[PAST 1ST SG.]	A&M[ti]-[PAST 1ST SG.]	ANCHORRσ
a. nafal	*!	
b. ☞ nafalti		*

In the following sections I turn to the analysis of hypocoristics and show that they adhere to the same constraint hierarchy argued for in the previous chapters, whereby the minimal word constraints dominate output-output faithfulness.

5.1. Serbo-Croatian

Serbo-Croatian has four main hypocoristic suffixes which are attached to truncated names.[20] The examples in (62) are representative:

(62) a. Feminine [-ka]
yélka < yélena
svétka < svétlana
b. Masculine [-ko]
mírko < míroslav
bránko < bránislav
zvónko < zvónimir
c. Masculine [-an]
míran < míroslav
vládan < vládimir
slóban < slóbodan
d. Masc. & Fem. [-a]
i. yéla < yélena
dúša < dúšanka
l^{y}íl^{y}a < l^{y}íl^{y}ana
ii. bóra < bórislav
vláda < vládimir
bóža < bóžidar

The truncated hypocoristic form to which one of the suffixes is attached is the first C_oVC of the base. This sequence may or may not correspond to a syllable in the base form. The syllabic as well as the morphological structure of the base form plays no role in the truncation process (e.g. *mí.ro+slav > mír.ko, mí.ran*, *l^{y}í.l^{y}a+na > l^{y}í.l^{y}a*).

[20] A fifth suffix, /-ica/, also exists, but it is the feminine diminutive suffix in the language and therefore the PrWd formed with this suffix does not adhere to the minimal word constraint (minimally and maximally di-syllabic/bimoraic foot). yélica < yélena, mírica < mírjana, etc. Note that the bare trancatum (i.e. the shortened form without the suffix) is the same as that in the other two types (**yél**ica, **yél**a, **yél**ka), possibly due to output-output correspondence.

As argued in previous chapters, truncation occurs when the minimal word constraints are ranked above faithfulness ones. Assuming the item-and-process model within OT, the suffixes are attached to the base via an alignment constraint (63) which aligns the suffix to the right (hence suffix) and matches its category (HYPO) to that required in the input.

(63) A&M[a] Align and Match ([a]-HYPO, R, PrWd-$_{[Base\text{-}HYPO]}$, R)
Align the right edge of the HYPO suffix [a] with the right edge of the PrWd including the base specified for HYPO, and match the category of the suffix with that required by the base.

The input of hypocoristics is the full surface name and the requirement that it be a hypocoristic, as shown in (64).

The A&M[a] constraint is satisfied even when part of the base is missing, as the alignment is to the PrWd containing the base (or part thereof).

The output hypocoristic must be no longer and no shorter than the minimal word in the language, so when a suffix is added, less material from the base can be present in the output.

In the following tableau, none of the candidates considered violate CONTIGUITY or LINEARITY and so the two constraints have been suppressed from the tableau, for space reasons.

(64)

Base: bórislav-$_{[HYPO]}$	MINWD	A&M[a]	ANCHORL I-O	ANCHORL O-I	MAXSEG
a. (bo.ri).(sla.va)]$_{PrWd}$	*!(ALLFTL)			*******	
b. (bo.ri).sla] $_{PrWd}$	*!(PARSE)			*****	2
c. (bo.ri).a] $_{PrWd}$	*!(PARSE)			***	4
d. ☞ (bo.ra)] $_{PrWd}$				**	5
e. (bo.ri).(slav)] $_{PrWd}$	*!(ALLFTL)	*		*******	
f. (bo.ris)] $_{PrWd}$		*!		****	3
g. (bo.ri)] $_{PrWd}$		*!		***	4

Candidate (a), though maximally faithful to the base and fulfills the HYPO requirement, fails to meet the MinWd requirement and is therefore eliminated from the competition. All other candidates violating any of the minimal word constraints (b, c, and e) are discarded as well. Candidates (f-g) are eliminated for violating the morphological constraint A&M[a], which is also undominated as the suffix is mandatory. The remaining candidate (d), which does not violate any of the undominated constraints, is thus selected as the optimal output.

Similar constraints to that in (63) must be posited for each suffix, as in (65). These constraints are not ranked with respect to each other, and so more than one output may emerge as optimal (tableau (66)).

(65) A&M[ka] Align and Match ([ka]-HYPO, R, PrWd-[Base-HYPO], R)
Align the right edge of the HYPO suffix [ka] with the right edge of the PrWd including the base specified for HYPO, and match the category of the suffix with that required by the base.

(66)

Base: yélena-[HYPO]	MINWD	A&M[a]	A&M[ka]	ANCHORL I-O	ANCHORL O-I	MAXSEG
a. ☞ (yé.la)]$_{PrWd}$			*		**	3
b. ☞ (yél.ka)] $_{PrWd}$		*			**	3
c. (yé.le)(na)] $_{PrWd}$	**!(FTBIN, ALLFTL)		*		*****	

When the vowel in the first syllable of the base is followed by a consonant cluster (e.g., *persida*) the longer output is selected (*persa*), as it least violates MAXSEG. However, this is true for the data in (67a), but not in (67b).

(67)

Base	Hypocoristic	
a. ándʒelka	ʔana	ándʒa
pérsida	*pera	pérsa
yézdimir	*yeza	yézda

b.	svétlana	svéta	*svetla
	rádmila	ráda	*radma
	mómčilo	móma	*momča

The two groups differ in the ranking of ANCHORL$_{O\text{-}I}$ above or below MAXSEG. The hypocorsitics in (67a) keep the consonantal cluster at the expense of more violation marks of ANCHORL$_{O\text{-}I}$ (68)) and in this respect they behave like the English CCs, while those in (b) prefer the shorter form at the expense of MAXSEG (tableau (69)) behaving in this respect as AWs. Notice that ándʒa (< ándʒelka) vs. móma (<mómčilo) suggest that σContact (Vennemann 1988, Clements 1990) does not play a role here, or that an additional constraint requiring the clustered consonants to have the same place of articulation may be necessary.

(68)

Base: pérsida -[HYPO]	MINWD	A&M[a]	ANCHORL$_{I\text{-}O}$	MAXSEG	ANCHORL$_{O\text{-}I}$
a. ☞ (per.sa)]$_{PrWd}$				2	**
b. (pe.ra)]$_{PrWd}$				3!	*

(69)

Base: svétlana -[HYPO]	MINWD	A&M[a]	ANCHORL$_{I\text{-}O}$	ANCHORL$_{O\text{-}I}$	MAXSEG
a. (svet.la)]$_{PrWd}$				***!	3
b. ☞ (sve.ta)]$_{PrWd}$				**	4

In the following section I discuss Hebrew hypocoristics, where I will show that the analysis discussed thus far holds for Hebrew as well, even though Hebrew allows two possible output forms.

5.2. Hebrew

Two types of Hebrew hypocoristics require the suffix [-i] to be added to either full names (70), or, more commonly, to truncated names (71). [21]

(70) Unmodified Hypocoristics

dorít > doríti

yafít > yafíti

odéd > odédi

ronén > ronéni

mixál > mixáli

revitál > revitáli

(71) Modified Hypocoristics

daniél > dáni

revitál > révi

mordexái > mórdi

odéd > ódi

xána > xáni

There are two conditions to forming **unmodified** hypocoristcis: a) the base name must end in a consonant (*dafna* > **dafnai, rivka* > **rivkai*); b) stress must fall on the last syllable (*orén* > **oréni, ayelét* > **ayéleti*). Any name that meets these two conditions may be added the suffix /-i/ as a term of endearment.[22]

The modified hypocoristics in (71) behave like the Serbo-Croatian ones in the previous section:

[21] The names are usually given in their normative form where stress is usually ultimate. The disyllabic names baring penultimate stress (*xána, malka*) relate to either normative *xaná, saráh, leá*, or to existing nouns baring ultimate stress (*malká* 'queen', *simxá* 'joy', *riná* 'singing').

[22] The suffix /-uš/, which is also used for endearment, has the same phonological distribution as /-i/, so that only names that meet the aforementioned conditions may be added the suffix /-uš/ (*dorít* > *dorítuš, odéd* > *odéduš,* but *orén* > **orénuš, dafna* > **dafnauš*).

(72)

Base: odéd-HYPO	MINWD	A&M[i]	ANCHORL I-O	ANCHORL O-I	MAXSEG
a. (ó.ded)		i!		***	
b. (ó.de).di	*!(PARSE)			***	
c. o.(dé.di)	*!*(PARSE, ALLFTL)			***	
d. ☞ (ó.di)				*	2

Candidates (b) and (c) are both faithful to the base, but are discarded for violating PARSEσ. Candidate (a) is entirely faithful to the base without violating the minimal word constraints, but violates the morphological constraint as it does not contain the suffix [-i]. Finally, candidate (d) is selected as the optimal output.

(73) Hypocoristics with internal clusters

i.	(y)irmiyáhu >	*(y)íri	(y)írmi
	avrahám >	ávi	ávri
	gavriél >	gábi	gávri
	málka >	máli	málki
ii.	nimród >	nímri	*nímri
	efrát/efráim >	éfi	*éfri

In case of consonantal clusters (73), sometimes the shorter output emerges (ii) and other times the longer one does too (i). Again, as in the Serbo-Croatian cases in (67), for the former group $\text{ANCHORL}_{\text{O-I}}$ is ranked above MAXSEG (as in AWs) and for the latter group, below MAXSEG (as in CCs), resulting in a greater variety of hypocoristics.[23]

The following section is concerned with English hypocoristics where the issue of faithfulness of the hypocoristic vowel is discussed.

[23] Although the argumet being made here is that hypocoristics are phonologically no different than clipped words, considerations outside the realm of phonology, such as pragmatics, may play a role.

5.3. English

5.3.1. General

English has two main forms of hypocoristics: one with no suffix (74a) and one with the suffix [-i] (74b).

(74)

Base		a. No Suffix	b. [-i]
samántha	>	sám	sámi
wíllliam	>	wíll	wílli
róbert	>	rób	róbi
déborah	>	déb	débi
éduard	>	éd	édi
jénifer	>	jén	jéni

The suffixed hypocoristics in (74b) result from the same basic constraint ranking as in Serbo-Croatian (section 5.1) and Hebrew (section 5.2), where the morphological constraints outrank the faithfulness ones. This basic hierarchy is responsible for the presence of the suffix. However, in many cases, as in (74), the suffixless hypocoristics exist in free variation with the suffixed forms. In OT, variation results from different rankings (Prince & Smolensky 1993). Thus, for the suffixless English hypocoristics, the morphological constraints are ranked below the faithfulness ones, as in (75):

(75)

Base: samantha	MinWd	ANCHORL$_{I\text{-}O}$	ANCHORL$_{O\text{-}I}$	MAXSEG	A&M[i]
a. ☞ (sam)			**	4	*
b. (sá.mi)			***!	4	
c. (sá.man)			***!*	2	*

ANCHORL$_{O\text{-}I}$ ranked above A&M[i] is responsible for selecting the suffixless output (candidate a) out of the candidates which survive the evaluation of the undominated constraints.

5.3.2. Vowel Faithfulness

An issue which has not been raised thus far concerns the indentity of the hypocoristic vowel to its correspondent in the base.

Benua's model of truncation discussed at the beginning of this chapter does not allow correspondence between input and truncated form. This predicts that the truncated form will never be more faithful to the input than the base. Benua shows that this accounts for the behaviour of English hypocoristics, summarised here for convenience.

In some dialects of English, the low front vowel [æ] does not appear before a tautosyllabic [r] in which case it is realised as [ɑ].

(76) English [æ] ~ [ɑ] Distribution

a.			b.		
	map	[mæp]		mar	[mɑr]
	carry	[kæ.ri]		car	[kɑr]
	Harry	[hæ.ri]		hard	[hɑrd]

The hypocoristics in (77), however, have [æ], not [ɑ], even though they precede a tautosyllabic [r].

(77) English Hypocoristics

Harry	[hæ.ri]	Har	[hær]
Larry	[læ.ri]	Lar	[lær]
Sarah	[sæ.ra]	Sar	[sær]

In this model, truncated forms, regulated by BT-Identity constraints, are faithful to their source, so that markedness constraints militating against tautosyllabic [ær] play no role. Also, in the dialects that drop the [r] in coda position, the [r] is not dropped in the truncated forms in (77), again, due to BT-Identity relations rather than IO-faithfulness.

In English, vowels in unstressed open syllables are typically reduced to schwas (Kahn 1976), as shown in (78).

(78) Vowel Reduction

Input	Output	Hypocoristic	
/æmændæ/	[ə.mǽn.də]	[mǽn.di] *[ǽ.m(i)]	Amanda
/rɛbɛkæ/	[rə.bɛ́.kə]	[bɛ́.ki] *[rɛ́b(i)]	Rebecka
/vɛrɔnikæ/	[və.rɔ́.nɪ.kə]	[rɔ́.ni] *[vɛ́r(i)]	Veronica
/mirændæ/	[mə.rǽn.də]	[rǽn.di] *[mír(i)]	Miranda
/pætriʃæ/	[pə.tʃrɪ́.ʃə]	[pǽt]	Patricia
/mɛlisa/	[mə.lí.sə]	[mɛ́l]	Melissa

A constraint militating against schwas in stressed position (*ə́) seems to be active here. Left-anchored hypocoristics cannot emerge from the base names in (78) which bear non-initial stress, as there would be no way to retrieve the underlying vowel under an output-output only correspondence model. One way of deriving hypocoristics from these forms is by violating ANCHORL$_{I\text{-}O}$. The data in (78) seems to provide evidence supporting Benua's two correspondence relations model. However, another way of deriving hypocoristics from names with an initial unstressed open syllable, is by keeping the underlying non-reduced vowel so as not to violate the *ə́ constraint and at the same time satisfying the ANCHORL$_{I\text{-}O}$ constraint, as in (79).

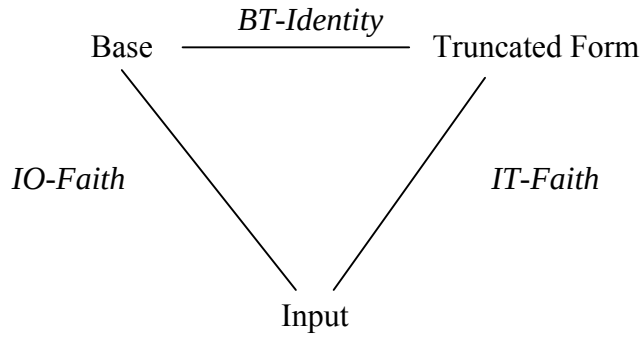
(79) Vowel Reduction with corresponding L-Anchored hypocoristics

Input	Ouput	Hypocoristic
/sæmænθæ/	[sə.mǽn.θə]	[sǽm]
/pɛnɛlɔpi/	[pə.nɛ́.lə.pi]	[pɛ́ni]
/mɛlisa/	[mə.lí.sə]	[mɛ́l]
/nɪkɔl/	[nə.kóʊl]	[nɪ́.ki]
/rɛjinæ/	[rə.jí.nə]	[rɛ́.ji]

Without any correspondence between input and truncated form, the vowels in the examples in (79) cannot be retrieved.

Benua's model, therefore, needs to be revised to allow correspondence between the truncated form and the input form:

(80) Revised Model



The relevant IT-Faith constraint is FAITHV_{IT}, which requires the output vowel to be identical to the underlying vowel. There is no evidence of crucial ranking between FAITHV_{IT} and its BT counterpart (FAITHV_{BT}), which requires the output vowel to be identical to the base vowel. However, ranking *ə́ above FAITHV_{BT} ensures that in case of a schwa in the base form, it is the input vowel that emerges (candidate (b)). Observe, that a candidate comprising a vowel other than the input vowel (candidate (c)) violates both FAITHV_{IT} and FAITHV_{BT} and is therefore discarded, even though it satisfies the *ə́ constraint.

(81)

Input: /sæmænθæ/ Base: [sə.mǽn.θə]	*ə́	FAITHV_{IT}	FAITHV_{BT}
a. (sə́m)	*!	*!	
b. ☞ (sǽm)			*
c. (sím)		*	*!

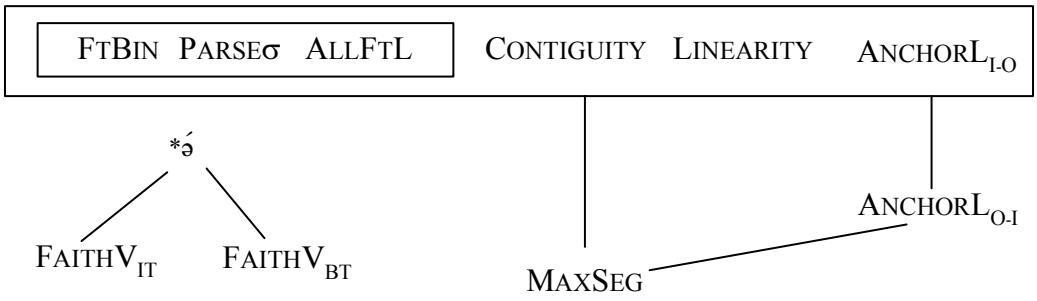
Hence, the hypocoristics reported by Benua (77) are not more faithful to the base than to the input, but rather they do not show the effect of the markedness constraint

against tautosyllabic [ær], which in the grammar of abbreviations is ranked below FAITHV_{BT}.

The forms in (78), which also display a schwa in the base's leftmost syllable, do not retrieve the underlying vowel in order to avoid a stressed schwa. Instead, they remain faithful to the base's stressed syllable, in violation of ANCHOR$L_{I\text{-}O}$. For these forms, then, the constraint requiring faithfulness to the stressed syllable of the base (MAX$_{\acute{\sigma}}$) outranks ANCHOR$L_{I\text{-}O}$.[24]

The following diagram summarises the hierarchy of the constraints argued for in this chapter.

(82) Constraint hierarchy



In the following chapter, I attempt to compare the abbreviation types discussed in this work, and to point out the similarities and the differences between the languages explored.

[24] This study is only concerned with left-anchored abbreviations, and so this point will not be further explored.

Chapter 6. A Comparative View

6.1. Introduction

The types of abbreviations discussed in this study have been shown to differ in the number of words comprising the base, and in the number of segments provided by each base word, as illustrated in the following table:

(83)

Abbreviation Type	No. of Base Words	No. of Segments from Each Base Word
AW	More than 1	1
CC		More than 1
Hypocoristics/Clippings	1	

The minimal word constraints have been shown to be responsible for the prosodic shape (a single binary foot) of all four types of abbreviation processes, subject to language specific foot inventory: AWs (Chapter 2), CCs (Chapter 3), Clippings (Chapter 4), and Hypocoristics (Chapter 5). The differences observed in the linguistic literature and throughout this study have been argued to result from the difference in the input forms rather than the output, and in some cases from a limited difference in the ranking of the constraints.

All of the abbreviation (shortening) processes generate a new prosodic word, a minimal word (which is also the maximal word) where the prosodic (markedness) constraints outrank faithfulness ones. However, the similarities between the types of abbreviation in question go beyond this generalisation. In fact, AWs and CCs are in complementary distribution and are therefore the same (section 6.2), Clippings are instances of CCs where the input is a single PrWd (section 6.3) and since hypocoristics are clippings involving personal names (section 6.4), it must be concluded that it is the input that affects the output shape and it is therefore unnecessary to assume separate grammars for these processes.

6.2. Acronym Words and Clipped Compounds

From a correspondence point of view, these two processes differ in the number of segments retained from each word in the base: if a single segment, the result is an AW, if more than one segment is retained from at least one base word, then the result is a CC (as in (83) above).

From an input point of view, the two processes may differ in two properties: (i) the type of segments residing at the left edge of the base words, where the initial segments alone are syllabifiable (84a), vs. where they are not (84b and 85a); or (ii) in the number of base words, only two (85b and c) vs. three or more (all the rest).

(84) a. Serbo-Croatian

sív < Savezno Izvršno Veće	'Federative Executive Assembly'
súp < Sekretaryat Unutrašnih Poslova	'Ministry of Internal Affairs'
ját < Jugoslovenski Aero Transport	'Yugoslav Air Transport'

b. Hebrew

ʔábax < **ʔ**atomi **B**iyologi **X**imi	'atomic biological & chemical' (warfare)
cáhal < **C**va **H**agana **L**e-israel	'Israeli Defence Force'
natbág < **N**emal **T**eufa **B**en-**G**urion	'Ben-Gurion airport'

c. English

ASCII (æski:) < American Standard Code for Information Interchange

NATO (néytoʊ) < North Atlantic Treaty Organisation

LASER (léyzə:) < Light Amplification (by) Stimulated Emission of Radiation

(85) a. Serbo-Croatian

EKOS < EKoloških Organizaciya Srbiye	'Ecology organisation of Serbia'
IGMA < Industriya Grajevinskog MAteriyala	'Industry of building material'
KOMGRAP < KOMunalno GRAjevinsko Preduzeće	'Community building company'

b. Hebrew

mankál < **MEN**ahel **KL**ali	'general manager'
ʔavác < **ʔAV**taxa **C**iburit	'public security'

c. English

sítcom < SITuation COMedy

ávgas < AViation GASoline

fórtran < FORmula TRANslation

Pure acronyms, as in (84a) and (84c), require vowels at the left edge of at least one base word for proper syllabification. Since Hebrew has (more frequently than not) consonants in these positions, the only pure AWs are those involving consonants that alternate with vowels in medial and final position. The relevant alternations between consonants and vowels in Hebrew are: y~i; v~o/u (see Bat-El 1994).

(86) y~i alternations

kis	'pocket'	kayas	'pick-pocket'
dira	'apartment'	dayar	'resident'
min	'type'	miyen	'sorted'

(87) v~o/u

dud	'boiler'	dvadim	'boilers'
šor	'bull'	švarim	'bulls'
dóar	'mail'	davar	'postman'
dox/duax	'report'	diveax	'reported 3rd masc.'

When these consonants are in initial position, they surface as consonants in the CC (88a). When in medial or final position, they surface as vowels (88b).

(88) a. Initial position

Base	CC	Gloss
vaada leteum miluim	valtam	'army reserves coordination committee'
vaadat xakira	vaxak	'commission of inquiry'
yexidat binuy	yaxab	'construction unit'

b. Medial and final position

ci yisraeli misxary	cim *cayam	'Israeli commercial fleet'
beyt israel lexu venelxa	bilu	a name of a pioneer group
yehuda vešomron	yoš	'Judea and Samaria'

Initial glides surface as consonants (88a) since the requirement that syllables have onsets (ONSET) is ranked above the constraint against glides, *GLIDE. This is illustrated in the following tableau.

(89) ONSET >> *GLIDE

yexidat binuy	ONSET	*GLIDE	*$V_{[\neq low]}$	IDENTF
a. ☞ yaxab		*		**
b. ixab	*!		*	*
c. yexib		*	*!*	

In *cim*, where the medial base word is vowel initial, the vowel is kept, even though this violates *$V_{[\neq low]}$, and the competing candidate **cayam* is discarded. *GLIDE is therefore ranked above *$V_{[\neq low]}$.

(90) *GLIDE >> *$V_{[\neq low]}$

ci israeli misxari	*GLIDE	MINWD	ANCHORL_{I-O}	*$V_{[\neq low]}$	IDENTF
a. ☞ cim		*(FTBIN)		*	
b. cayam	*!				*
c. cam		*(FTBIN)	*!		*

From an output point of view, however, a view supported by OT, these differences dissolve. Both AWs and CCs have the same prosodic shape: a single foot.

When the first segments alone can be properly syllabified, AWs may surface, but if their syllabification is impossible, either due to the number of base words or to

the type of segments at the left edge of the base words, then additional segments are extracted and CCs emerge. AWs and CCs are, therefore, in complementary distribution, as illustrated in tableaux (91) (where an AW emerges) and (92) (where a CC emerges).

(91)

north atlantic treaty organisation	MINWD	ANCHORL$_{I-O}$	ANCHORL$_{O-I}$	MAXSEG
a. ☞ (na.to)				26
b. (na.torg)			*!*	22

(92)

situation comedy	MINWD	ANCHORL$_{I-O}$	MAXSEG	ANCHORL$_{O-I}$
a. (sc)	*!$_{(FTBIN)}$		13	
b. ☞ (sit.com)			9	****

In section 3.1 I proposed to rank the two ANCHORL constraints such that ANCHORL$_{I-O}$ dominates ANCHORL$_{O-I}$. This accounted for the fact that the former constraint is never violated in any of the abbreviation types while the latter is frequently violated (except in AWs). Furthermore, for English CCs, ANCHORL$_{O-I}$ is ranked below MAXSEG. Serbo-Croatian and Hebrew CCS have the same constraint ranking as that of AWs.

The basic ranking argued for is as follows:

(93) AW and CC Basic Constraint Ranking
MINWD, LINEARITY, CONTIGUITY, ANCHORL$_{I-O}$ >> ANCHORL$_{O-I}$ >> MAXSEG

The following lists the language specific diversions from this basic ranking:

(94)

	AW	CC
Serbo-Croatian	FAITHLENGTH >> FTBIN, FTSAL, TROCHQU (MinWd) >> DEP-H (for monosyllabic outputs)	MAXSEG >> ALIGNCODA (final codas are preferred to medial codas)
Hebrew	*GLIDE >> MINWD >> *$V_{[\neq LOW]}$ (for pure AW, where vowels other than *a* emerge)	*$V_{[\neq LOW]}$ >> IDENTF (accounts for the vowel *a*)
		FINALC >> ANCHORL_{O-I} (avoids CvCv outputs in 2-words bases)
English		MAXSEG >> ANCHORL_{O-I} (selects the longest di-syllabic output)

6.3. Clipped Compounds and Clippings

Clippings are instances of CCs, where the base comprises a single PrWd. Therefore, in order to comply with the minimal word requirement, all of the segmental material must be taken from this PrWd.

(95) Clippings

a. Hebrew

sú**per** < **super**market	'supermarket'
trí**go** < **trigo**nometria	'trigonometry'

b. English

exám < **exam**ination

ad < **ad**vertisement

mémo < **mem**orandum

The two sets of data in (95) above differ from CCs in the number of base words, but result from the same constraints, though the few Hebrew clippings that exist retain the base vowels (96), while CCs do not (97).

(96)

supermarket	MINWD	ANCHORL$_{I-O}$	FINALC	ANCHORL$_{O-I}$	MAXSEG
a. (sup)	*!(FTBIN)			**	8
b. (su.pe)			*!	***	7
c. ☞ (su.per)				****	6

(97)

yexidat binuy	MINWD	ANCHORL$_{I-O}$	FINALC	ANCHORL$_{O-I}$	*V[≠low]	MAXSEG
a. (yab)	*!(FTBIN)			*		14
b. (ya.ba)			*!	**		13
c. ☞ (ya.xab)				***		13

Other Hebrew Clippings end in a suffix, as discussed in the following section.

English clippings have been shown to have the same grammar of AWs in that ANCHORL$_{O-I}$ outranks MAXSEG, and that some clips require a disyllabic output. As proposed in Chapter 4, English has two types of clippings: one where the required foot is moraic, and another where it is syllabic.

The basic ranking for CCs and clippings is as in (93). The following lists the language specific diversions from this basic ranking:

(98)

	CC	Clippings
Hebrew	*V$_{[\neq LOW]}$ >> IDENTF	IDENTF >> *V$_{[\neq LOW]}$
	FINALC >> ANCHORL$_{O-I}$	FINALC >> ANCHORL$_{O-I}$
English	MAXSEG >> ANCHORL$_{O-I}$	FTBINσ >> FTBINμ (group 2)

6.4. Clippings and Hypocoristics

Some Hebrew Clippings (as in (99)) end in the suffix /-i/.

(99) Hebrew Clippings

díḱi < dikaon	'depression'
ríḱši < rigšot (ashma)	'feeling of guilt'
kárci < karciya	'tick - (for someone who is nagging)'

Through this suffix, the Hebrew clips acquire a (sometimes derogatory) hypocoristic-like quality completely absent from the suffixless ones in (95).

Likewise, English too has clipped forms with a suffix (/-o/) and as a result have the same (derogatory) hypocoristic-like quality to them:

(100) psýcho < **psycho**path
schízo < **schizo**phrenic
fatso < **fat**
weirdo < **weird**

As argued in Chapter 5, the hypocoristic is associated with a suffix via an A&M$_{[SUFF]}$ constraint, which aligns the suffix at the right edge and matches it to the requirement coded in the input. This morphological constraint is necessarily ranked above ANCHORL$_{O-I}$, as the suffixed segments do not have correspondents in the base.

For the suffixless hypocoristics, which adhere to the same set of constraints argued for clippings, these morphological constraints (one for each suffix) have been argued to be ranked below ANCHORL$_{O-I}$, allowing it to select the shortest suffixless candidate.

An additional issue raised in the discussion of hypocoristics, relates to the correspondence relations between input (UR) and the abbreviated form (i.e. IT-Identity). One group of English hypocoristics, in which the first syllable of the base (which is an output form) contains an unstressed reduced vowel, provided evidence that contrary to Benua's claim, correspondence between input and the truncated form must be allowed. Without this correspondence relation, the underlying vowel could not be retrieved for left-anchored hypocoristics, resulting in impossible forms with a stressed schwa.

Assuming the same basic constraint hierarchy as in (93), the following lists the devaitions from it in clippings and hypocoristics:

(101)

	Clippings	Hypocoristics	
		With Suffix	Suffixless
Serbo-Croatian		A&M$_{[SUFF]}$ >> ANCHORL$_{O-I}$	
		For some: MAXSEG >> ANCHORL$_{O-I}$	
Hebrew	FINALC >> ANCHORL$_{O-I}$	A&M$_{[SUFF]}$ >> ANCHORL$_{OI}$	ANCHORL$_{O-I}$ >> A&M$_{[SUFF]}$
		For some: MAXSEG >> ANCHORL$_{O-I}$	
English	FTBINσ >> FTBINμ (group 2)	A&M$_{[SUFF]}$ >> ANCHORL$_{O-I}$	ANCHORL$_{O-I}$ >> A&M$_{[SUFF]}$
		FAITHV$_{IT}$ >> FAITHV$_{BT}$	

6.5. Concluding Remarks

In this study I aimed to show the similarities between four abbreviation processes. In general, as illustrated in (83), (i) AWs emerge from bases with more than one word, and only one segment is taken from the left edge of each base word, and the more words in the base, the higher the chance of obtaining an AW; (ii) CCs emerge from bases with more than one word, and more than one segment is taken from at least one base word; and (iii) hypocoristics and clippings emerge from single-word bases, from which more than one segment forms the abbreviation. I claimed that the prosodic constraints ranked above faithfulness constraints trigger abbreviation and that the basic constraints responsible for the selection of the base segments are identical for all four processes (see (93)).

It should be noted, however, that some differences are found between the languages, as well as between processes. I have pointed out in Chapter 2 that while it is required that the output abbreviation be a foot, the languages differ in the type of

feet allowed: (i) in Hebrew, where weight plays no role in the language's phonology, foot status can be reached only syllabically (FTBIN$_\sigma$); (ii) English, being sensitive to weight, allows feet to be monosyllabic, if bimoraic (FTBIN$_\mu$). However, some clippings require that the foot be di-syllabic, suggesting for them the ranking FTBIN$_\sigma$ >> FTBIN$_\mu$; Serbo-Croatian also allows monosyllabic feet, if attached to a high tone (FTSAL).

For the discussion of tonal feet, FAITHLENGTH was introduced to explain why no disyllabic/bimoraic tonal feet are allowed; and since the acronym bases do not contain tone, DEP-H needed to be introduced as well.

In Hebrew, (i) *GLIDE is ranked above the MinWd constraints to account for the AWs including a medial vowel, which appear to be subminimal; (ii) FINALC, ranked above ANCHORL$_{O\text{-}I}$ was necessary to account for CCs with two base-words and for clippings which avoid CvCv outputs; (iii) *V$_{[\neq LOW]}$ is responsible for the vowel *a* which emerges in CCs instead of the vowel inherent in the base.

Finally, (i) in English CCs, ANCHORL$_{O\text{-}I}$ is ranked below MAXSEG, as the optimal outputs are not the shortest allowed foot (this is also process-specific); (ii) *ə́ >> FAITHV$_{IT}$, FAITHV$_{BT}$ is responsible for retrieving the underlying vowels of hypocoristics whose bases contain a schwa. This ranking, however, does not conflict with the ranking of the other constraints in any of the languages discussed.

Furthermore, while it is argued that the constraints involved in all of the abbreviation processes discussed in this study are the same, some constraints do seem to be process-specific. While the vowels in Hebrew CCs are invariably *a*, in Clippings and Hypocoristics the base vowels are preserved, so that the processes differ in the degree of markedness. Also, a difference in the ranking of ANCHORL$_{O\text{-}I}$ above or below MAXSEG, allows for shorter or longer outputs respectively, while

adhering to the minimal word requirement. This was seen in some Hebrew and Serbo-Croatian hypocoristics as well as in English CCs.

Finally, in all three languages, the morphological constraints $A\&M_{[SUFF]}$ accounted for the hypocoristic suffix when ranked above $ANCHORL_{O\text{-}I}$.

The following table summarises all of the differences between the languages and between the various abbreviation types, excluding the morphological constraints which are ranked above $ANCHORL_{O\text{-}I}$ only in abbreviations involving suffixes (namely hypocoristics and some clippings).

	AW	CC	Clippings	Hypocoristics
Serbo-Croatian	FAITHLENGTH >> FTBIN, FTSAL, >> DEP-H	MAXSEG>>ALIGNCODA		For some: MAXSEG>>$ANCHORL_{O\text{-}I}$
Hebrew	*GLIDE >> MINWD	FINALC>>$ANCHORL_{O\text{-}I}$ $*V_{[\neq LOW]}$>>IDENTF		For some: MAXSEG>>$ANCHORL_{O\text{-}I}$
English		MAXSEG>>$ANCHORL_{O\text{-}I}$	FTBINσ>>FTBINμ	

Appendix A. Hebrew Data

A.1. AWs and CCs

	Base	AW/CC	Gloss.
1.	agaf koax adam department personnel	áka	Department of personnel
2.	agaf modiin department intelligence	amán	Department of intelligence
3.	anašim xašuvim meod people important (pl.) very	axám	VIP
4.	atomi biyologi ximi atomic biological chemical	ábax	Atomic biological chemical warfare
5.	beyt israel lexu venelxa home israel go and we go (future)	bílu	Name of pioneer group
6.	ci israeli misxari fleet israeli commercial	cim	Israeli commercial fleet
7.	ciyud bdika equipment test	cabád	Testing equipment
8.	cva hagana le-israel army defence for israel	cáhal	Israeli Defence Force
9.	drišat šalom request hello	daš	Greetings
10.	ecem bilti mezohe object unidentified	abám	UFO
11.	irgun (le)-šixrur palestina organisation (for) liberation palestine	ášaf	PLO
12.	karov labait close to house	kaláb	Close to home
13.	madrix kluim instructor prisoners	madáx	Prisoners instructor
14.	madrix sport instructor sport	madás	Sports instructor
15.	mate klali headquarters general	matkál	General headquarters
16.	mazkir klali secretary general	mazkál	Secretary general
17.	mefakeax klali inspector general	mafkál	Inspector general
18.	mefaked basis commander base	mabás	Base commander
19.	mefaked kita commander class	mak	Class commander

	Base	AW/CC	Gloss.
20.	mefaked xativa commander brigade	maxát	Brigade commander
21.	mefaked xeyl (ha)-yam commander corps (the) sea	máxi	Navy commander
22.	menahel klali manager general	mankál	CEO
23.	merkaz maxševim (ve) rišum centre computers (and) record memukan automated	mamrám	Automated Computer Centre
24.	mexkar u-pituax research and development	mop	Research and development
25.	mifkedet xativa headquarters brigade	mafxát	Brigade headquarters
26.	mifkedet xeylot sade headquarters corps (pl.) field	mafxáš	Infantry headquarters
27.	mišmar (ha)-gvul guard the border	magáv	Border guard
28.	mispar katalogi number catalogue (adj.)	makát	Catalogue number
29.	nemal teufa Ben-Gurion port flight Ben-Gurion	natbág	Ben-Gurion airport
30.	peilut xablanit oyenet activity terrorism hostile	páxa	Hostile terrorist activity
31.	pikadon kcar moed deposit short time	pakám	Short term deposit
32.	plugot maxac troops wound	palmáx	Shock troops
33.	praklit cvai raši attorney military chief	pacár	Chief military attorney
34.	rav-samal baxir sergeant senior	rasáb	Senior sergeant
35.	roš liška head office	raláš	Head of office
36.	segen mišne lieutenant secondary	sagám	Second lieutenant
37.	šerut bitaxon klali service security general	šabák	Israeli secret service
38.	sgan aluf deputy champion	saál	Lieutenant colonel
39.	sgan mefaked pluga deputy commander company	samáp	Company deputy commander

	Base	AW/CC	Gloss.
40.	šituf peula cooperation activity	šatáp	Collaboration
41.	šlita u-bakara command and control	šob	Command and control
42.	ʔumot meuxadot nations united (fem. pl.)	ʔum	United Nations
43.	vaada xakira committee investigation	vaxák	Investigation committee
44.	xaver kneset member Kneset (Israeli parliament)	xak	Member of parliament
45.	xavurat pikud kidmi group command front	xapák	Front command group
46.	xayal pašut soldier simple	xapáš	Private
47.	xomer lexima ximi material combat chemical	xaláx	Chemical warfare
48.	yehuda ve-šomron Judea and Samara	yoš	Judea and Samara
49.	yexidat binuy unit of construction	yaxáb	Construction unit

A.2. Clippings and Hypocoristics

	Base	Hypocoristic
1.	rivká	rivi, riki
2.	dan, daniél	dáni
3.	yonatán	yóni
4.	mixaél	míki
5.	menaxém	méni, xémi
6.	binyamín	bíni, bíbi
7.	revitál	révi
8.	yoséf	yósi, séfi
9.	asáf	ási, sáfi
10.	eliyáhu	éli
11.	avrahám	ávi, ávri
12.	yafít	fíti

	Base	**Hypocoristic**	
13.	tamár	támi	
14.	yaakóv	yáki, kóbi	
15.	cipóra	cípi	
16.	yirmiyáhu	yírmi	
17.	šim'ón	šími	
18.	mixál	míxi, míki	
19.	miriám	míri	
20.	símxa	símі, síma	
21.	gavriél	gábi, gávri	
22.	málka	máli	
23.	efráim	éfi	
24.	dafná	dáfi	
25.	nimród	nímі	
26.	odéd	ódi	
27.	bircinút	bírci	‘seriously’
28.	súpermarket	súper	‘supermarket’
29.	trigonométria	trígo	‘trigonometry’
30.	dikaón	díki	‘depression’
31.	rigšot (ašma)	rígši	‘feeling of guilt’
32.	karciyá	kárci	‘tick’

Appendix B. English Data

B.1. AWs

	Base Elements	Acronym	AW
1.	Absent Without Official Leave	AWOL	éɪwɒl
2.	Acquired Immune Deficiency Syndrome	AIDS	éɪdz
3.	Advisory Council On Science and Technology	ACOST	əkɒ́st
4.	Airborne Warning And Control System	AWACS	éɪwæks
5.	American Standard Code for Information Interchange	ASCII	ǽskiː
6.	Anti Submarine Detection Investigation Committee	ASDIC	ǽzdɪk
7.	Association of Scientific, Technology, and Managerial Staffs	ASTMS	ǽztɛmz
8.	Australian and New Zealand Army Corps	ANZAC	ǽnzæk
9.	Boot And Shoe Manufacturer's Association	BASMA	bǽsmə
10.	British Association of Social Workers	BASW	bǽzwə
11.	Computer Aided Design	CAD	kæd
12.	Congress Of Racial Equality	CORE	kɔː
13.	Cost Of Living Adjustment	COLA	kə́ʊlə
14.	English for Speakers of Other Languages	ESOL	íːsɒl
15.	Federal Emergency Management Agency	FEMA	fĭmə
16.	Global Atmospheric Research Programme	GARP	gɑːp
17.	Indeterminate Mass Particle	IMP	ɪmp
18.	Jet-Assisted Take-Off	JATO	ǰéɪtəʊ
19.	Judge Advocate General	JAG	ǰæg

	Base Elements	Acronym	AW
20.	Light Amplification by Stimulated Emission of Radiation	LASER	lé͡ɪzə
21.	Low Altitude Navigation and Targeting Infra-Red for Night	LANTIRN	lǽ͡ntən
22.	National Aero-Space Agency	NASA	nǽ͡sə
23.	National Emergency Alarm Repeater	NEAR	niə:
24.	North Atlantic Treaty Organisation	NATO	né͡ɪtəʊ
25.	Programmed Logic for Automatic Teaching Operation	PLATO	plé͡ɪtəʊ
26.	Quiet Take-Off and Landing	Q-TOL	kyú͡:tɒl
27.	Random Access Memory	RAM	ræm
28.	Read Only Memory	ROM	rɒm
29.	Self Contained Underwater Breathing Apparatus	SCUBA	skú͡:bə
30.	Special Weapons And Tactics	SWAT	swɒt
31.	Surface-to-Air Missile	SAM	sæm
32.	United Nations Educational, Scientific, and Cultural Organisation	UNESCO	yú͡:nɛskəʊ
33.	United Nations Interim Force In Lebanon	UNIFIL	yú͡:nɪfɪl
34.	United Nations International Children's Emergency Fund	UNICEF	yú͡:nɪsɛf
35.	Visual Average Speed Computer And Recorder	VASCAR	vǽ͡skɑ:
36.	What You See Is What You Get	WYSIWYG	wí͡zɪwɪg
37.	White Anglo Saxon Protestant	WASP	wɒsp
38.	Wings, Engine, Fuselage, Tail	WEFT	wɛft
39.	Women's Environmental Network	WEN	wɛn

	Base Elements	Acronym	AW
40.	Young, Attractive, Verbal, Intelligent, and Successful	YAVIS	yǽvɪs
41.	Youth International Party	YIP	yɪp

B.2. CCs

	Base	CC	
1.	advanced decoy technology	adtech	ǽdtɛk
2.	airways communication (system)	aircom	ɛ́ːkɒm
3.	algorithmic language	algol	ǽlgɒl
4.	amphibious tractor	amtrac	ǽmc̆ræk
5.	aviation gas	avgas	avgas
6.	Because It's Time network	bitnet	bɪ́tnɛt
7.	Belgium, Neterhlands, Luxemburg	Benelux	bɛ́nəlʌks
8.	common business oriented language	cobol	kə́ʊbɒl
9.	communications satellite	comsat	kɒ́msæt
10.	computer service	compuserve	kɒ́mpyusəːv
11.	conference student organization of linguistics in Europe	console	kɒ́nsəʊl
12.	formula translation	fortran	fɒ́ːtrən
13.	high fidelity	hi-fi	haɪ-faɪ
14.	international police	interpol	ɪ́ntəpɒl
15.	long range navigation	loran	lə́ʊræn
16.	mass concentration	mascon	mǽskɒn
17.	miniature camera	minicam	mɪ́nɪkæm
18.	national communications	natcom	nǽtkɒm

	Base	CC	
19.	parallax second	parasec	pǽrəsɛk
20.	picture element	pixel	píksəl
21.	radio detecting and ranging	radar	réɪdɑː
22.	reduction oxidation	redox	rɛ́dɒks
23.	science fiction	sci-fi	saɪ-faɪ
24.	situation comedy	sitcom	sítkɒm
25.	sound navigation ranging	sonar	sə́ʊnɑː
26.	will comply	wilco	wílkəʊ
27.	windows information	winfo	wínfəʊ
28.	windows magazine	winmag	wínmæg

B.3. Clippings

	Base	Clip
1.	advertisement	ad, advert
2.	bachelor	bach
3.	champion	champ
4.	condominium	condo
5.	discotheque	disco
6.	examination	exam
7.	fraternity	frat
8.	graduate	grad
9.	gymnaesium	gym
10.	homosexual	homo
11.	luncheon	lunch
12.	mathematics	math

	Base	Clip
13.	memorandum	memo
14.	metropolitan	metro
15.	nitroglycerine	nitro
16.	photograph	photo
17.	preliminary	prelim
18.	preparatory	prep
19.	proffesional	pro
20.	psychology	psych
21.	puppy	pup
22.	recreation	rec
23.	rhinoceros	rhino
24.	sister	sis
25.	synchronise	sync
26.	trigonometry	trig
27.	veterinarian	vet
28.	wizard	wiz
29.	yankee	yank
30.	kilogram	kilo
31.	hippopotamus	hippo

B.4. Hypocoristics

	Base	Hypocoristic	
1.	bɛ́njəmɪn	bɛ́n(i), bɛ́nj(i)	‘Benjamin’
2.	dɛ́ɪvɪd	dɛ́ɪv(i)	‘David’
3.	nɪ́kəʊləs	nɪ́k(i)	‘Nicholas’

	Base	**Hypocoristic**	
4.	maɪkl̩	mík(i), máɪk(i)	‘Michael’
5.	jɛ́nɪfə:	jɛ́n(i)	‘Jenifer’
6.	jɔ́nəθən	jɔ́n(i)	‘Johnathan’
7.	wílyəm	wíl(i), bíl(i)	‘William’
8.	rʊ́bə:t	rʊ́b(i), bʊ́b(i)	‘Robert’
9.	ɑ́:θə:	ɑ́:t(i)	‘Arthur’
10.	dǽnɪəl	dǽn(i)	‘Daniel’
11.	gɛ́ɪbrɪəl	gɛ́ɪb, gǽbi	‘Gabriel’
12.	rítʃə:d	rík(i), rítʃ(i), dík	‘Richard’
13.	ɛ́duə:d	ɛ́d(i)	‘Eduard’
14.	dɛ́bra	dɛ́b(i)	‘Deborah’
15.	səmǽnθə	sǽm(i)	‘Samantha’
16.	sǽmyʊəl	sǽm(i)	‘Samuel’
17.	kǽθrɪn	kǽt, kǽθi, kɛ́ɪt(i)	‘Catherine’
18.	jóʊzəf	jóʊ(i)	‘Joseph’
19.	mɑ:grət	mæg(i)	‘Margaret’
20.	tʃɑ́:lz	tʃɑ:li	‘Charles’
21.	frɛ́drɪk	frɛ́d(i)	‘Frederick’
22.	ɑ́:nəld	ɑ́:ni	‘Arnold’
23.	ǽlfrəd	ǽlf(i)	‘Alfred’
24.	pətʃrɪ́ʃə	pǽ́t(i)	‘Patricia’

Appendix C. Serbo-Croatian Data

C.1. AWs

	Base Elements	AW	Gloss.
1.	akademsko kulturno umetnićko društvo	ákud	'Academic cultural artistic company'
2.	antifašistićko veče narodnog oslobajenya yugoslaviye	ávnoy	anti-fascistic council of national liberation of Yugoslavia
3.	apatinska modna obuća	ámo	'Apatin Footware Fashion'
4.	auto motor sport klub	ámsk	'Automobile sports club'
5.	autonomna kossovsko metohijska oblast	ákmo	'Autonomous Kossovo and Metohija Area'
6.	beogradska revija amaterskih malih scena	bráms	'small amateur theatre organisation of Belgrade'
7.	elektro-distribuciya beograd	édi	'Belgrade electric distribution'
8.	fabrika armatura, specijalnih mašina i alata	fásma	'Factory for fossets, special machinery and tools'
9.	narodno oslobodilačka borba	nób	'Battle of national liberation'
10.	narodno oslobodilaćki rat	nór	'War of national liberation'
11.	nedel^{y}ne informativne novine	nín	'Weekly informative newspaper'
12.	penzionersko amatersko kolturno umetnićko društvo	pákud	'retirement artistic and cultural company'
13.	savezno izvršno veće	sív	'Federative executive assembly'
14.	sekretaryat unutrašnih poslova	súp	'Ministry of internal affairs'

	Base Elements	AW	Gloss.
15.	yugoslovenski aero transport	yát	'Yugoslav air-transport'

C.2. CCs

	Base	CC	Gloss.
1.	beogradska konfekciya	béko	'Belgrade confection'
2.	ekoloških organizaciya srbiye	ékos	'Ecology organisation of Serbia'
3.	fabrika motora sarayevo	fámos	'Sarajevo motor plant'
4.	gradski magazin	grámag	'City store'
5.	grajevinski metal	grámet	'Metal industry'
6.	gradska narodna prodavnica	gránap	'City community store'
7.	gradski tek͡stil	gráte͡ks	'City textile'
8.	hemiyski proizvodi	hémpro	'Chemical products'
9.	industriya grajevinskog materiyala	ígma	'Industry of building material'
10.	komunalno grajevinsko preduzeće	kómgrap	'Community building company'
11.	narodni magazin	náma	'national magazine'

C.3. Hypocoristics

	Base	Hypocoristic
1.	ánjelka	ánja, ána, ánka
2.	bórislav	bóra
3.	bórislava	bórka
4.	bóžidar	bóža, bóško, bóžan
5.	bránislav/bránimir	bráne, bránko

	Base	**Hypocoristic**
6.	bránislava	brána, bránka
7.	čédomir	čéda
8.	dánka	dána
9.	dávor	dárko
10.	désanka	désa, déska
11.	drágoslav	drági, drágan
12.	dúšan	dúško
13.	dúšanka	dúša, dúška
14.	kátarina	káta, káća
15.	kséniya	séna, kséna, sénka
16.	l^{y}íl^{y}ana	l^{y}íl^{y}a
17.	míilica	míila, míilka
18.	míodrag	míya, míyan
19.	míroslav	míro, míirko, míran
20.	míryana	míra, míirka
21.	mómčilo	móma
22.	nádežda	náda
23.	nátalya	náta, nátka
24.	nédel^{y}ko	néja
25.	pérsida	pérsa, pérka
26.	rádmila	ráda
27.	slávomir	sláva, slávko
28.	slobodánka	slóba

	Base	**Hypocoristic**
29.	slóbodan	slóba. slóban
30.	smíl^{y}ana	smíl^{y}a, smíl^{y}ka
31.	snéžana	snéža, snéžka
32.	srbol^{y}ub	srba, srban
33.	svétlana	svéta, svétka
34.	svétozar	svéta, svétko, svétan
35.	tómislav	tóma
36.	vásil^{y}e	vása, vásko
37.	vélimir	vél^{y}a, vél^{y}ko, vél^{y}an
38.	vésela	vésa, véska
39.	vládimir	vláda, vládan
40.	yélena	yéla, yélka
41.	yézdimir	yézda, yézdan
42.	yóvan	yóva
43.	yóvanka	yóka, yóvka
44.	želimir	žél^{y}a, žél^{y}ko, žél^{y}an
45.	živoyin	živa, živko, živan
46.	zvónimir	zvóne, zvónko

References

Acronyms Dictionary: First Edition. (1960). Detroit: Gale.

Adam, Galit. (to appear). *From Variable to Optimal Grammar: Evidence from Language Acquisition and Language Change*. Doctoral Dissertation, Tel-Aviv University.

Adam, Galit, and Outi Bat-El. (2000). "Morphological Knowledge without Morphological Structure: The Emergence of Inflectional Suffixes in the Acquisition of Hebrew Verbs." A paper presented in a workshop on Phonology and Prosodic Morphology of Semitic Languages. Tel-Aviv.

Anderson, Stephen R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark. (1976). *Word Formation in Generative Grammar*. Cambridge, MA: The MIT Press.

Bat-El, Outi. (1994). "The Optimal Acronym Word in Hebrew." In P. Koskinen (ed.) *Proceedings of the 1995 Annual Conference of the Canadian Linguistic Association*. Toronto: University of Toronto, 23-38.

Bat-El, Outi. (1999). "To Chop or Not to Chop and What to Chop: Colloquial Hebrew Truncated Imperatives." A paper presented at the colloquia in Tel-Aviv University and the Hebrew University. Tel-Aviv and Jerusalem.

Bat-El, Outi. (2000a). "The Grammaticality of 'Extragrammatical' Morphology." In Doleschal and Thornton (eds.) *Extragrammatical and Marginal Morphology*. Munich: Lincom Europa, 61-84.

Bat-El, Outi. (2000b). "Anti-faithfulness: An Inherent Morphological Property." A paper presented at the 5th International Conference of Afroasiatic Languages. Paris.

Bauer, Laurie. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.

Bauer, Laurie. (1988). *Introducing Linguistic Morphology* 1994 ed. Vol. Edinburgh: Edinburgh University Press.

Benua, Laura. (1995). "Identity Effects in Morphological Truncation." In J. Beckman, S. Urbanczyk and L. Walsh (eds.) *Papers in Optimality Theory: University of Massachusetts Occasional Papers in Linguistics*. Vol. 18. Amherst, Massachusetts: GLSA, 77-136.

Benua, Laura. (1997). *Transderivational Identity: Phonological Relations Between Words*. Doctoral Dissertation, UMASS, Amherst. ROA-259.

Bolotsky, Shmuel. (1982). "Remarks on Rhythmic Stress in Modern Hebrew". *Journal of Linguistics*. 18: 275-289.

Cannon, Garland. (1989). "Abbreviations and Acronyms in English Word-Formation." *American Speech* 64.2: 99-127.

Clements, G. N. (1990). "The Role of the Sonority Cycle in Core Syllabification." In Kengston and Beckman (eds.). *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*, 283-333. Cambridge: Cambridge University Press.

Clements, G. N., and S. Jay Keyser. (1983). *CV Phonology: A Generative Theory of the Syllable*. Cambridge MA: MIT Press.

Davis, Stuart, and Bushra Adnan Zawaydeh. (1999). "A descriptive analysis of hypocoristics in Colloquial Arabic." *Languages and Linguistics* 3: 83-98.

Dixon, Robert M. W. (1972). *The Dyribal Language of North Queensland*. Cambridge: CUP.

Dixon, Robert M. W. (1977). "Some Phonological Rules in Yidiny." *Linguistic Inquiry* 8: 1-34.

Dressler, Wolfgang U. and Annmarie Karpf. (1994). "The Theoretical Relevance of Pre- and Protomorphology in Language Acquisition." In Booij and Van Marle (eds.) *Yearbook of Morphology 1994*. Dordrecht: Kluwer Academic Publications, 99-122.

Graf, Dafna. (2000). "Stress Assignment in the Nominal System of Modern Hebrew (MH)." In Adam Z. Wyner (ed.) *The Proceedings of the 15th Annual Conference of IATL 2000*. Vol. 7

Hammond, Michael. (1995). "There Is No Lexicon." Ms., University of Arizona, Tucson.

Hammond, Michael. (1997). "Optimality Theory and Prosody." In Diana Archangeli and D. Terence Langendoen (eds.) *Optimality Theory: An Overview*. Reprinted 1998. Malden: Blackwell, 33-58.

Hayes, Bruce. (1987). "A Revised Parametric Metrical Theory." In J. McDonough and B. Plunkett (eds.), *Proceedings of NELS 17*. Amherst, MA: GLSA, 274-289.

Hayes, Bruce. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Hockett, Charles. (1954). "Two Models of Grammatical Description." *Word* 10:210-231.

Itô, Junko, and Armin Mester. (1997). "Sympathy Theory and German Truncations." In Viola Miglio and Bruce Morén (eds.) *Proceedings of the Hopkins Optimality Workshop/Maryland Mayfest 1997*. Vol. 5: University of Maryland Working Papers in Linguistics.

Kager, René. (1989). *A Metrical Theory of Stress and Destressing in English and Dutch.* Dordrecht: Foris.

Kager, René. (1991). "The Moraic Iamb." In L. Dobrin, L. Nichols, and R. Rodrigues (eds.), *Papers from CLS 27*. Chicago: CLS, 291-306.

Kahn, Daniel. (1976). *Syllable-Based Generalizations in English Phonology*. Doctoral Dissertation, MIT. New York: Garland Press, 1980.

Kennedy, Arthur G. (1935). *Current English*. Boston: Ginn.

Lieber, Rochelle. (1992). *Deconstructing Morphology*. Chicago: The University of Chicago Press.

McCarthy, John J. (1993). "A Case of Surface Constraint Violation." *Canadian Journal of Linguistics* 38:169-195.

McCarthy, John J., and Alan S. Prince. (1986). *Prosodic Morphology*. Ms., University of Massachusetts, Amherst, and Brandeis University, Waaltham, Mass.

McCarthy, John J., and Alan S. Prince. (1994). "The Emergence of the Unmarked." In Mercè Gonzàles (ed.) *Proceedings of NELS 24*. Amherst, MA: Graduate Linguistic Student Association, 339-379.

McCarthy, John J., and Alan S. Prince. (1995). "Faithfulness and Reduplicative Identity." *Papers in Optimality Theory: University of Massachussetts Occasional Papers*. Vol. 18. Amherst Mass.: Graduate Linguistic Student Association., 249-384.

Mester, R. Armin. (1990). "Patterns of Truncation." *Linguistic Inquiry* 21(3): 478-485.

Piñeros, Carlos-Eduardo. (1999) *Prosodic and Segmental Unmarkedness in Spanish Truncation*. (ROA 332). 44 p.

Poser, William. (1990). "Evidence for Foot Structure in Japanese." *Lg* 66: 78-105.

Prince, Alan S. (1980). "A Metrical Theory for Estonian Quantity." *Linguistic Inquiry* 11: 511-562.

Prince, Alan S., and Paul Smolensky. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms, Rutgers University and University of Colorado, Boulder.

Russell, Kevin. (1995). "Morphemes and Candidates in Optimality Theory." *ROA-44-0195*.

Russell, Kevin. (1999). "MOT: Sketch of an OT Approach to Morphology." *ROA-352-1099*.

Selkirk, Elisabeth. (1980a). "Prosodic Domains in Phonology: Sanskrit Revisited." In Mark Aronoff and Mary-Louise Kean (eds.) *Juncture*. Saratoga, CA: Anma Libri, 107-129.

Selkirk, Elisabeth. (1980b). "The Role of Prosodic Categories in English Word Stress." *LI* 11: 563-605.

Selkirk, Elisabeth. (1983). *The Syntax of Words*. Cambridge, MA: The MIT Press.

Smolensky, Paul. (1996). "The Initial State and 'Richness of the Base' in Optimality Theory. ROA-154.

Vennemann, Theo. (1988). *Preference Laws for Syllable Structure.* Berlin: Mouton de Gruyter.

Weeda, Donald. (1992). *Word Truncation in Prosodic Morphology*. Doctoral dissertation, Univresity of Texas, Austin.

Yip, Moira. (1998). "Identity Avoidance in Phonology and Morphology." In S.G. Lapointe, D.K. Brentari and P.M. Farrell (eds). *Morphology and Its Relation to Phonology and Syntax*. Stanford: CSLI Publications. Pp.216-246. Also *ROA-82*.

Zec, Draga. (1999). "Footed Tones and Tonal Feet: Rhythmic Constituency in a Pitch-Accent Language." *Phonology* 16: 225-264.