

Inflectional Morphology in Optimality Theory

A Dissertation Presented

by

Zheng Xu

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

August 2007

Copyright by
Zheng Xu
August 2007

Stony Brook University

The Graduate School

Zheng Xu

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Mark Aronoff, Professor of Linguistics

Robert D. Hoberman, Professor of Linguistics

Alice C. Harris, Professor of Linguistics

**James P. Blevins, Assistant Director of Research
Research Center for English and Applied Linguistics
University of Cambridge**

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Inflectional Morphology in Optimality Theory

by

Zheng Xu

Doctor of Philosophy

in

Linguistics

Stony Brook University

2007

This dissertation proposes an inferential-realizational model of inflectional morphology (Matthews 1972, Zwicky 1985, Anderson 1992, Aronoff 1994, Stump 2001) within the framework of Optimality Theory (Prince and Smolensky 1993). Following Russell 1995, Yip 1998, Hyman 2003, MacBride 2004, I assume that the phonological information of inflectional affixes is introduced through realization constraints (RC) which associate abstract morphosyntactic or semantic feature values with phonological forms. I propose that rankings of realization constraints conform to the specificity condition, i.e. a constraint realizing a more specific morphosyntactic feature value set outranks a less specific realization constraint. I also propose that the unmarked situation in which one feature value is realized by one form (Wurzel 1989) is encoded in two universal and violable markedness constraints, *FEATURE SPLIT which bans the realization of a feature value by more than one form and *FEATURE FUSION which bans a form realizing more than one feature value.

Based on this model, I examine language phenomena such as OCP-triggered selection of phonologically unrelated (allo)morphs in Greek, Hungarian, Tswana, and Spanish, ordering of inflectional affixes in Lezgian, blocking of inflectional affixes

and extended morphological exponence in languages like Tamazight Berber, and directional syncretism in languages like Latin.

I show that this model has advantages over other morphological models in several ways. (1) It readily captures cases in which a default marker emerges to replace a morphosyntactically more specific marker which is expected to be adjacent to a phonologically similar form (OCP >> RC_{specific} >> RC_{less specific}). By contrast, the relation between a more specific marker and a less specific one needs to be stipulated in the input in a model which introduces phonological information through inputs (e.g. Bonet 2004). (2) It readily captures universal generalizations on affix order (Greenberg 1963, Bybee 1985), e.g. a number exponent cannot be farther away from a nominal stem than a case exponent because case scopes over number. Such generalizations are missed in Paradigm Function Morphology (Stump 2001) without extraordinary machinery. (3) Based on rankings of *FEATURE SPLIT and constraints realizing the same morphosyntactic feature value(s), it provides a unified account of both blocking and extended exponence without recourse to either a distinction between primary and secondary exponents (Noyer 1992) or multiple rule blocks (Stump 2001). (4) Based on output-to-output correspondence constraints (Benua 1995, McCarthy and Prince 1995), it readily captures cases of divergent bidirectional syncretism (Baerman 2004) in which syncretism brings about both marked and unmarked forms, a problem for Noyer 1998, which claims that syncretism always moves from a more marked to a less marked state.

Table of Contents

Acknowledgments.....	viii
1 Introduction.....	1
1.1 Questions.....	1
1.2 A realization OT model.....	3
1.3 Other theoretical frameworks.....	6
2 A realization OT approach to (avoidance of) repetition of identical morphs.....	8
2.1 Introduction.....	8
2.2 Two types of Optimality-Theoretic approaches.....	10
2.2.1 Mandarin <i>le</i>	10
2.2.2 A morphologically restricted OT account of Mandarin <i>le</i> haplology.....	11
2.2.3 A realization OT account.....	12
2.2.4 Cases of repetition of identical morphs.....	13
2.3 Problems for a morphologically restricted OT model.....	16
2.4 A realization OT approach to avoidance of repetition of (partially) identical morphs.....	19
2.4.1 Greek negatives.....	19
2.4.2 Hungarian second person singular indefinite.....	21
2.4.3 Tswana predicative concords.....	25
2.4.4 Spanish clitics.....	30
2.4.5 Swedish haplology.....	35
2.5 Problems for realization approaches and their solutions.....	37
2.6 Conclusion.....	40
3 A realization OT approach to affix order in Lezgian.....	46
3.1 Introduction.....	46
3.2 Lezgian.....	46
3.2.1 Nominal inflection.....	46
3.2.2 Verbal inflection.....	55
3.3 A realization OT approach to affix order in Lezgian.....	61
3.3.1 Lezgian nominal inflection.....	63
3.3.2 Lezgian verbal inflection.....	65
3.4 Other approaches.....	71
3.5 Conclusion.....	73
4 A realization OT approach to blocking and extended morphological exponence.....	78
4.1 Introduction.....	78
4.2 *FEATURE SPLIT and Pāṇini's Principle.....	79

4.3 A realization OT approach to blocking and extended exponence.....	84
4.3.1 Tamazight Berber.....	84
4.3.2 Classical Arabic.....	95
4.3.3 More examples of extended exponence.....	100
4.4 Alternative approaches to blocking and extended exponence.....	102
4.5 Conclusion.....	102
5 A realization OT approach to full and partial identity of forms.....	106
5.1 Introduction.....	106
5.2 Paradigmatic syncretism.....	106
5.2.1 Divergent bidirectional syncretism.....	108
5.2.2 A realization OT account of paradigmatic syncretism.....	111
5.3 Partial identity of forms.....	116
5.3.1 Pinker 1998.....	116
5.3.2 A realization OT account of partial identity of forms.....	117
5.3.3 Rules of referral.....	119
5.4 Conclusion.....	119
6 Conclusion.....	122
References.....	126

Acknowledgments

It is very difficult to write this section because I have benefited from so many people during the past few years and it would probably take another dissertation to acknowledge each of them in detail. I am very grateful to all those who have helped me with my life and study at Stony Brook University. From the bottom of my heart, I would like to say “thank you!”

Following the tradition, I would like to first acknowledge my dissertation committee members. First and foremost, I would like to thank my advisor and mentor, Mark Aronoff. I first met Mark at the department party to welcome newcomers when I came to Stony Brook. He came to me and asked whether I was interested in working as his editorial assistant for *Language*. I happily accepted the job and our relation was established then. He introduced me to morphology, a fascinating field to which I will probably devote the rest of my life. Every meeting with him was a joy. Mark had tons of jokes and interesting stories though he never joked about work. He was a great scholar and had an amazingly rich knowledge of language and linguistics. From him I learned not only how to do linguistics but also how to interact with people. He was always nice to me and always had time for me even though for some time he had a heavy administrative burden as a deputy provost of Stony Brook University. He read every abstract and paper I wrote, and always made highly insightful comments on them. He supported me both spiritually and financially from the moment I came to Stony Brook University until the moment I left. Without his guidance and encouragement, I would never have reached this stage. My greatest luck at Stony Brook was to have Mark as my advisor. I could never thank Mark enough for his long-time support and fatherly care.

I would like to thank Alice Harris, another important figure in my career. Alice came to Stony Brook University several years after me. She taught Historical Linguistics, Morphology, and a doctoral seminar on ergativity all of which I attended. Because of her wonderful personality and rich knowledge of language and linguistics, I decided to work with Alice. She found many chances for me so that I could become more professional. Alice was very popular with graduate students. She was always nice and always talked with a smile though she was very serious about academics and would not let pass a small mistake in a paper. I found it great fun to work with Alice.

I would also like to thank Bob Hoberman, another dissertation committee member. Bob and I had a long-standing friendship. He was also on the committee of my morphology qualifying paper. Almost every meeting with Bob lasted several hours because he could generate so many interesting ideas, questions, and comments. Bob was a nice, gentle, and quiet scholar but he was straightforward on what he disliked. I relied on Bob's knowledge of Semitic linguistics to flesh out the chapter on blocking and extended morphological exponence. It was real fun to work with Bob.

I also wish to thank Jim Blevins, my outside committee member. I met Jim and his family at Alice's house. Jim was very much into inflectional morphology and naturally became my outside member. Jim wrote very critical comments on my dissertation proposal so that my dissertation was very much improved. He also raised

very good questions at the dissertation defense which could lead to several follow-up projects. Jim, thanks so much for being on my committee!

I would also like to thank the faculty members at the Department of Linguistics at Stony Brook University. I thank Frank Anshen for his time on making annual tax forms for us. Mark, Frank, and I co-authored a paper on deponency in Latin. I thank Ellen Broselow whom most of my knowledge of phonology came from. I learned from her not only phonology but also devotion to academics and dedication to students. Special thanks go to Richard Larson who was very interested in East Asian languages. He was the best lecturer I had ever seen. His teaching was always clear, well-prepared, and word-perfect. He gave me lots of encouraging suggestions in various respects. In addition, I thank John Bailyn who taught me Historical Syntax, Christina Bethin who taught me Phonology I, and Dan Finer who taught me several syntax courses. Dan always raised very interesting questions and made very helpful comments on our work. I also thank Marie Huffman and Lori Repetti for whom I worked as a teaching assistant. Their knowledge of their fields highly impressed me.

Special thanks go to my friends from the Department of Linguistics at Stony Brook University. In particular, I wish to thank the following three. Jon MacDonald was one of my best friends. We had a lot of discussions about linguistics and life in general. I sincerely hope that Jon and his wife Jessica will live a happy life. Lanko Marušič has been an important figure at the department not only because of his talent in linguistics but also because of so much work he did for the department including keeping the grad lab workable, designing homepages, and providing various kinds of technical support. Lanko and I had lots of debate about linguistics. He patiently listened to most of my academic ideas and made very critical comments on them. I thank Lanko for what he has done for me. Mark Volpe was a special friend of mine. He was also into morphology and we had many wonderful discussions about linguistics and life in general. His suggestions were always helpful and encouraging. We look forward to future good times. I sincerely hope that he will remain happy and healthy and live a wonderful life with his daughter Tami. Thanks also to Diane Abraham, Marianne Borroff, Yiya Chen, Carlos de Cuba, Susana Huidobro, Jiwon Hwang, Tomoko Kawamura, Ruiqin Miao, Anne Miller, Katharina Schuhmann, Chih-hsiang Shu, Irina Tarabac, Julie Weisenberg, and Hiroko Yamakido for making my experience at Stony Brook unique. I wish all of them great success in their future career. I also thank our department staff members Sandra and Susan who have addressed many of my questions and come to my aid many times in the past few years. I wish to give very special thanks to Harriet and Perry who made me feel at home.

Special thanks also go to my non-linguist friends either in the United States or in China, who have helped or supported me in various respects. These friends include Cao Jiong, Guan Yuhua, Guo Aimin, Li Zhen, Li Zongyuan, Luo Yi, Qiu Wei, Wu Qi, Yang Yan, Zhang Wei, Zhang Wenyi, Zhao Wei, and Zou Jingyu.

Finally, my deepest thanks to my family in China for their love and understanding. Without their support, I could never have reached this stage. With all my love, from the bottom of my heart, thank you!

Chapter One

Introduction

1.1 Questions

Joan Bybee (1985: 207) says that “[t]he study of language is the study of the relation of meaning to form. The search for a linguistic theory is the search for the language-specific and universal principles that govern this relation.” The relation of meaning to form plays a central role in many language phenomena. Greenberg 1963 found, for example, that “the expression of number almost always comes between the noun base and the expression of case” (Greenberg 1963: 112, cited in Bybee 1985: 34). See the following Finnish example in which the exponent of plural, *-i* is closer to the root *talo* than the exponent of inessive case, *-ssa*.

- (1) *talo-i-ssa-ni*
house-PLURAL-INESSIVE-1SG
‘in my house’ (Spencer 2003: 630)

The question is how to encode this language universal in a formal theoretical framework to explain similar language phenomena.

Additionally, in many languages a morphosyntactically less specific form often occurs to replace a morphosyntactically more specific form which is supposed to be adjacent to another form with a similar phonological shape. For example, in Hungarian the suffix *-sz* [s] is an exponent of the feature value set {second person, singular, indefinite, present tense, and indicative mood}. See (2a). By contrast, the suffix *-ol* which expresses a less specific feature value set {second person, singular, and indefinite} has a much wider distribution and basically occurs in tenses and moods other than the present indicative. See (2b) (in which the vowel [o] is subject to vowel harmony). Interestingly, the suffix *-ol* replaces the suffix *-sz* in the context of second person, singular, indefinite, present tense, and indicative mood when *-sz* is expected to be adjacent to a nominal stem ending in a strident. See (2c).

- (2) The distribution of *-sz* and *-ol* in Hungarian (Vago 1980: 50)
- | | | |
|-------------------------------------|--------------|-----------------|
| a. <i>vár-sz</i> | ‘wait for’ | (present tense) |
| b. <i>vár-t-ál</i> | ‘waited for’ | (past tense) |
| c. <i>hoz-ol</i> (* <i>hoz-sz</i>) | ‘bring’ | (present tense) |

The question is what kind of formalism can capture this cross-linguistic phenomenon which involves the alternation of both a morphosyntactically more specific and less specific form.

Moreover, inflectional affixes, which may belong to different positions in an inflected word, often compete to realize a morphosyntactic or semantic feature value. Consider the following Classical Arabic paradigm. The prefix *t-* is a marker of either second person or third person feminine. The prefix *y-* is a marker of third person. The suffix *-na* is a marker of feminine plural and realizes both second and third persons feminine plural. Interestingly, in the paradigmatic slot of third person feminine plural, *y-* shows up instead of *t-*. The empirical generalization as made by Noyer 1997 is that both *-na* and *t-* express feminine and *-na* blocks the occurrence of *t-*. Moreover, in the slot of second person feminine singular, the prefix *t-* which expresses second person co-occurs with the suffix *-ii* which expresses second person, feminine, and singular.

(3) Partial Classical Arabic verbal paradigm in subjunctive mood (adapted from Noyer 1997: 5)

singular	dual	plural	
<i>t-aktub-a</i>	<i>t-aktub-aa</i>	<i>t-aktub-uu</i>	2, masc
<i>t-aktub-ii</i>	<i>t-aktub-aa</i>	<i>t-aktub-na</i>	2, fem
<i>y-aktub-a</i>	<i>y-aktub-aa</i>	<i>y-aktub-uu</i>	3, masc
<i>t-aktub-a</i>	<i>t-aktub-aa</i>	<i>y-aktub-na</i>	3, fem

The question is how to account for the cases in which *-na* {fem, plural} blocks the occurrence of *t-* {3, fem} while *t-* {2} co-occurs with *-ii* {2, fem, sg} given that both *-na* and *t-* express feminine and both *t-* and *-ii* express second person.

Apart from the above-mentioned phenomena, languages with an inflectional system sometimes exhibit *directional syncretism* (Stump 2001, Baerman 2004, Baerman, Brown, and Corbett 2004). That is, several paradigmatic cells share the same form and apparently the form of one cell is copied by others. Consider the following Latin declension. In the Latin second declension, the nominative singular copies the form of the accusative singular in the environment of default neuter nouns while the accusative singular copies the form of the nominative singular in the environment of a group of neuter nouns including *vulgus* ‘crowd’, *vi:rus* ‘poison’, and *pelagus* ‘sea’.

(4) The Latin second declension (adapted from Baerman 2004: 816)

	DEFAULT NEUTER ‘war’	DEFAULT MASCULINE ‘slave’	NOM & ACC in <i>-us</i> ‘crowd’
NOM SG	<i>bell-um</i>	<i>serv-us</i>	<i>vulg-us</i>
ACC SG	<i>bell-um</i>	<i>serv-um</i>	<i>vulg-us</i>

The question is how to describe directional syncretism within a theoretical model which also accounts for other cross-linguistic phenomena.

Following Bybee 1985, this dissertation studies “the relation of meaning to form” and searches for “the language-specific and universal principles that govern this relation.” It aims to develop a morphological model which can readily account for

language phenomena in which the relation of meaning to form plays a crucial role. Our focus is on inflectional morphology.¹ Specifically, I propose an inferential-realizational model of inflectional morphology (Matthews 1972, Zwicky 1985, Anderson 1992, Aronoff 1994, Stump 2001) within the framework of Optimality Theory (OT) (Prince and Smolensky 1993). I show that this model readily captures the above-mentioned language phenomena, which have been widely discussed in the literature.

1.2 A realization OT model

In this section I explain relevant terminologies and lay out fundamental assumptions within the proposed model. This model is inferential in that I assume an affix is *not* a lexical entry by itself but is instead introduced by a grammatical function, either a rule or a constraint (cf. Aronoff 1976, Zwicky 1985, Anderson 1992, Stump 2001). This model is realizational in that I assume affixation is *licensed* by abstract morphosyntactic or semantic feature values.

Following Russell 1995, Yip 1998, Hyman 2003, MacBride 2004, I assume that the phonological information of inflectional affixes is introduced through realization constraints (RC) which associate abstract morphosyntactic or semantic feature values with phonological forms.² For example, in Classical Arabic the prefix *y-* is an exponent of third person. We can encode the relation of the third person feature value to the prefix *y-* in a realization constraint (5). The realization constraint in (5) says that the third person feature value is realized (or spelled out) by the prefix *y-*. In a realization constraint, a morphosyntactic or semantic feature value is placed on the left of the colon while a phonological form is on the right. The colon can be read as “(to be) realized by.”

(5) {3}: *y-*

Additionally, I propose that rankings of realization constraints strictly obey the specificity condition stemming from Pāṇini’s Principle which was discovered by the Sanskrit grammarian, Pāṇini. That is, a more specific morphosyntactic feature value set has priority to be realized, so a constraint which realizes a set of feature values always outranks another constraint which realizes a subset of these feature values. For example, in Classical Arabic the constraint which realizes third person feminine by *t-* outranks the one which realizes third person by *y-*. See the ranking in (6).

(6) {3, fem}: *t-* >> {3}: *y-*

Notice that each realization constraint is a language-particular instantiation of a universal schema which associates meaning with form. See Mohanan and Mohanan 2003 which argues that we need to relax the OT assumption that all constraints are universal (Prince and Smolensky 1993). Mohanan and Mohanan 2003 argues for a model of generating language-particular constraints based on universal schemata. I assume that a universal schema in which a specific RC outranks a less specific RC

may also generate language-particular instantiations since realization constraints are language-particular.

In Optimality Theory, an input provides necessary information for the constraint-based grammar to process while an output is assumed to be produced by the function *Gen* which can generate an infinite list of logical possibilities. An optimal output candidate is the “winning candidate” selected by the grammar from all logical possibilities. Following Yip 1998, Hyman 2003, MacBride 2004, I assume that in the proposed realization OT model, an input consists of a lexeme (noun, verb, adjective) with its lexical information and abstract morphosyntactic or semantic feature values. An output consists of phonological information produced by the function *Gen*, which realizes both a lexeme and abstract feature values.

Additionally, I assume that non-phonological information such as morphosyntactic or semantic feature value, lexical category, diacritic feature remains identical in both the input and output. This assumption is somewhat in the spirit of Grimshaw (1997a) which presents an OT framework in syntax. Grimshaw (1997a) assumes that competing candidates to be generated for a single input must be semantically equivalent. This assumption is incompatible with a mechanism which changes non-phonological information such as feature-impoverishment (Noyer 1992, 1997, 1998), generation of morphosyntactic feature values (Grimshaw 1997b, 2001, Wunderlich 2001). Lumsden 1992 criticizes such feature-changing mechanisms and remarks that:

If these rules change the feature-values of the underlying positions, then the underlying distribution of features has very little relation to the semantic/syntactic environment. That is, there is no longer any systematic relationship between the distribution of features in syntactic positions and the semantic/syntactic significance that is associated with feature labels. Ultimately, this complex of rules has no motivation other than to mechanically describe the distribution of forms in the surface structure. (Lumsden 1992: 472-3, cited in Noyer 1997: 86)

Noyer (1997: 87) admits that “such [feature-changing] rules are highly costly. If alternative analyses exist, they are presumably less costly.”

Several assumptions also need to be made. Above all, since we assume that the phonological information of inflectional affixes is introduced through realization constraints, what then of lexemes? Is a Root or a lexical stem introduced through a constraint or input? Theories diverge on this point. In a non-constraint-based framework like Lexeme-Morpheme Base Morphology (Beard 1995), lexemes and inflectional affixes are treated differently. Beard (1995: 44) says that:

[L]EXEMES in LMBM are defined unexceptionally in terms of open class signs ... They are directly associations of properly specified sequences of

phonemes, grammatical features, and semantic intensions, that is, noun, verb, and adjective stems ... Bound grammatical morphemes, on the other hand, are defined as morphological spelling operations in the literal sense of 'morphological': modifications of the phonological form (Greek *morphé*) only of lexemes. These modifications MARK, EXPRESS, or SPELL the same closed grammatical categories, lexical and syntactic, as do free grammatical morphemes.

According to LMBM, it seems that we might want to specify the phonological information of a lexeme in the input so that we can distinguish lexemes from inflectional affixes whose phonological information is introduced through rules or constraints. Similarly, Stump 2001 treats inflectional affixes and lexemes differently. Inflectional affixes are introduced through realization rules while lexical stems are basically introduced through stem-selection rules. By contrast, Distributed Morphology (Halle and Marantz 1993) assumes that phonological information is introduced by *Late Insertion* so that both Roots and Vocabulary Items are inserted into syntactic terminal nodes after all syntactic processes. In this spirit, it seems that we might want to introduce the phonological information of lexemes through constraints since both Roots and Vocabulary Items undergo morphological spelling operations. Within a realization OT model, whether the phonological information of a lexeme is introduced through inputs or constraints depends on the author. For example, Russell 1995 assumes that the input does not contain any phonological information and therefore, no underlying forms. By contrast, Yip 1998, Hyman 2003, and MacBride 2004 assume that the phonological information of a lexeme is introduced through inputs. Following Beard 1995, Yip 1998, Hyman 2003, and MacBride 2004, I assume that the phonological information of a lexeme is introduced through an input.

Moreover, Optimality Theory (Prince and Smolensky 1993) relies on faithfulness constraints to maintain the information that occurs in the input. For example, the constraint DEP I-O requires no occurrence of information that does not exist in the input. The constraint MAX I-O requires no deletion of information that exists in the input. Faithfulness constraints are also incorporated into the proposed realization OT model. Following Yip 1998, I assume that an affix that is introduced by a realization constraint and occurs in the output violates the constraint DEP I-O. DEP I-O needs to rank lower than realization constraints because otherwise abstract morphosyntactic or semantic feature values would not be spelled out. Since it is assumed that the phonological information of a lexeme appears in the input, we need the constraint MAX I-O to ban any loss of input phonological information. I will further discuss these faithfulness constraints in accounting for cases of directional syncretism in which the form of a feature value set is copied by another set. See Chapter 5.

Markedness constraints lie in the heart of Optimality Theory. In the proposed realization OT model which deals with the relation of meaning to form, I propose that the ideal unmarked situation in which one meaning corresponds to one form (Wurzel 1989) is encoded in two universal and violable constraints *FEATURE SPLIT and *FEATURE FUSION which are defined as follows.

- (7) a. *FEATURE SPLIT: A morphosyntactic or semantic feature value cannot be realized by more than one phonological form.
- b. *FEATURE FUSION: A phonological form cannot realize more than one morphosyntactic or semantic feature value.

This ideal unmarked situation is what characterizes agglutinative languages. But many languages are not agglutinative, which suggests the violability of the markedness constraints favoring this unmarked situation and lends support for an OT model, which relies on the violability of constraints. Additionally, I propose that universal generalizations on affix order (Greenberg 1963, Bybee 1985) can be encoded in scope constraints which require, for example, a number exponent cannot be farther away from the same stem than a case exponent.

Furthermore, a realization constraint like {3}: *y*- can be decomposed into two constraints, one realizing the morph *y* and the other placing it before the Root. To describe the position of *y*, we can refer to an alignment constraint (McCarthy and Prince 1993, Russell 1997, Grimshaw 2001) which places an object in a surface position. For the sake of brevity and simplicity, I use the notation like {3}: *y*- and discuss alignment constraints when their occurrence is necessary.

1.3 Other theoretical frameworks

One of the major tasks of this dissertation is to show that the proposed realization OT model has advantages over other theoretical frameworks in accounting for language phenomena such as those in section 1.1.

I advocate a realizational model because as pointed out by Stump 2001, it is compatible with cases of extended exponence in which a morphosyntactic or semantic feature value is realized by more than one form. For example, in Classical Arabic the second person feature value is realized by both *t*- {2} and *-ii* {2, fem, sg}. By contrast, incremental models (Lieber 1992, Steele 1995, Wunderlich 1996) assume that affixation introduces morphosyntactic feature values. These models also assume that affixation is strictly information-adding and therefore must exclude cases like the one in Classical Arabic given that *-ii* introduces the second person feature value which does not need to be introduced by *t*- again.

I propose a constraint-based model because it readily encodes language universals in constraint-rankings. This model is output-oriented in that the grammar consists of restrictions on the output. I show that the proposed model readily captures cases like the one in Hungarian (2) in which a morphosyntactically less specific form emerges to replace a morphosyntactically more specific form which is supposed to be adjacent to another phonologically (partially) identical form. By contrast, it will be hard for non-constraint-based frameworks like Distributed Morphology (Halle and Marantz 1993) and Paradigm Function Morphology (Stump 2001) to capture, for example, the Hungarian case in (2) because these frameworks lack an inherent mechanism to force the less specific form to emerge under a powerful phonological restriction banning repetition of adjacent (partially) identical morphs. Additionally,

universal generalizations on affix order (Greenberg 1963, Bybee 1985) are missed in a framework like Paradigm Function Morphology which assumes that affix order is determined by the order of rule blocks on a language-particular basis.

Within Optimality Theory, there are two parallel morphological models. One assumes that the phonological information of affixes is introduced through inputs and morphological information such as “affix”, “root”, and “stem” constitutes enough information for the grammar to process (Prince and Smolensky 1993, McCarthy and Prince 1993b, Bonet 2004, among many others). I show that this model needs to make bald stipulations to deal with cases in which a morphosyntactically less specific form emerges to replace a morphosyntactically more specific form which is supposed to be adjacent to another phonologically (partially) identical form. It would miss a unified grammar to deal with, for example, the Classical Arabic case in which the second person feature value can be realized twice while the feminine feature value cannot. In contrast to this morphologically restricted OT model which is not centered on the relation of meaning to form, I show that the proposed realization OT model readily captures these phenomena in which the relation of meaning to form plays a crucial role. On the other hand, a model based on the generation of morphosyntactic or semantic feature values (Grimshaw 1997, 2001, Wunderlich 2001) is incapable of accounting for the language phenomena discussed in section 1.1 without extraordinary machinery because this model lacks an inherent system to spell out abstract feature values.

This dissertation investigates an autonomous morphological domain which is built on the relation of meaning to form. In contrast to a morphologically restricted OT model, the proposed realization OT model assumes that realization is a central part of an autonomous morphological domain, although phonological constraints sometimes show their effects in this domain. For example, allomorph selection is sometimes motivated by phonotactics. In contrast to a model which revels in every instance of syntax-morphology interpenetration, I assume that affix order is motivated primarily by semantic considerations, which make syntactic structure and morphological structure “mirror” each other (cf. Wunderlich and Fabri 1996).

Notes

¹ I do not aim to give an accurate definition of inflection which will be used to test tricky cases that can be argued to be derivational. See, for example, Wurzel 1989, Anderson 1992 for reviews of various criteria which fail to give a one-hundred-percent accurate definition of inflection.

² Yip 1998 argues that clitics, particles, and possibly function words are also introduced through realization constraints in addition to inflectional affixes.

Chapter Two

A realization OT approach to (avoidance of) repetition of identical morphs

2.1 Introduction

It has been observed that languages either ban or allow repetition of adjacent identical morphs (Stemberger 1981, Menn and MacWhinney 1984). Strategies to avoid such repetition include haplology, allomorph selection, and replacement by a different structure.

Stemberger (1981: 791) defines morphological haplology as a phenomenon in which “an affix or clitic is absent when the adjacent part of the stem is homophonous to it.” Examples can be easily found in English (1). When the English possessive marker *-s* is attached to the plural form of the noun *boy*, i.e., *boys*, we get *boys’* instead of the illicit form **boys’s*. (The symbol *** indicates ungrammaticality.)

(1) *boys* (plural) + *-s* (possessive) → *boys’* (**boys’s*)

Stemberger further remarks that morphological haplology is a universal phenomenon observed in both Indo-European and non-Indo-European languages.

Menn and MacWhinney 1984 give many examples of avoidance of repetition of adjacent identical morphs by (allo)morph selection (or suppletion). An example from Spanish is given in (2). In Spanish, when the 3rd person dative object clitic *le* precedes the 3rd person accusative object clitic *lo*, the spurious *se* rule (Perlmutter 1971) applies, i.e., the surface combination is *se lo* instead of the expected **le lo*.

(2) *le* (3rd person dative object) + *lo* (3rd person accusative object) → *se lo* (not **le lo*)

Moreover, Menn and MacWhinney 1984 point out that repetition of adjacent identical morphs can be avoided by using a different structure. For example, many English speakers prefer not to add the adverb-forming suffix *-ly* to adjectives ending in *-ly* (3). Instead, they either use synonyms or avoid sentences in which these illicit adverbs may occur.¹

(3) Adjective	Adverb
<i>manly</i>	<i>*manlily</i>
<i>likely</i>	<i>*likelily</i>
<i>ugly</i>	<i>*uglily</i>

Meanwhile, languages also allow repetition of adjacent identical morphs, as pointed out by Menn and MacWhinney. For example, in Choctaw the active voice suffix /-li/ can be adjacent to the first person singular (1sg) suffix /-li/.

A question naturally arises. How do we account for the above phenomena or how do we encode the above phenomena in grammars? Stemberger 1981 argues against the approach to morphological haplology which requires a rule to delete an identical morph to satisfy the surface structural requirement that two identical morphs should not be adjacent. He says that “the deletion rule creates a surface form identical to what would have been produced with no rules at all; thus no evidence exists for the application of either [a] rule [adding an identical morph or one deleting it then]” (p.803). Instead, Stemberger advocates an approach to haplology based on vacuous application of rules, i.e., if an input already satisfies the surface structural pattern, a rule which is supposed to change an input applies vacuously and leaves the input intact. This idea stems from Hooper’s 1976 proposal that “all phonological rules should be generalizations which are true of surface forms.” (I cite Stemberger’s (1981: 805) interpretation.) Thus, to account for cases of English haplology like *boys’* (**boys’s*), a rule which attaches the genitive marker -s to *boys* (plural) applies vacuously to satisfy both the structural requirements that a plural or genitive form end in -s and no identical morphs be adjacent.

In comparison, Menn and MacWhinney 1984 propose an affix-checking mechanism. Under this approach, when speakers attempt to attach, for example, the English genitive marker -s, they will check whether the base contains an affix-like -s or not. If it does, then the item which already contains an -s morph will block the attachment of the genitive -s. Meanwhile, they claim that there exists an output constraint banning two adjacent identical morphs or a morph adjacent to a stem whose proper subpart has an identical shape to the morph.

Previous works to explain the phenomena in (1-3) have noticed that avoidance of adjacent identical forms points toward an output-oriented model, i.e., a model based on restrictions on surface forms. Given this observation, cases of (avoidance of) repetition of identical morphs become grist for the mill of Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1993b), an output-oriented theory which encodes typological variation in constraint rankings.

This chapter is devoted to comparing two types of Optimality-Theoretic (OT) approaches to (avoidance of) repetition of identical morphs. In one, a morphologically restricted OT model assumes that morphological information such as “affix”, “root”, and “stem”, and phonological information related to phonetic/phonological features, segments, and suprasegmental properties constitute enough input and output information for the grammar to process. This model has traditionally been termed “prosodic morphology”. On the other hand, another OT approach, which is much less orthodox, assumes that the input contains unrealized morphosyntactic or semantic information such as case, number, etc. and the grammar contains a group of constraints associating morphosyntactic or semantic feature values with phonological forms. I call this approach a realization OT approach.² In this chapter, it is argued that a realization OT approach shows its edge over a morphologically restricted OT

approach in accounting for cases of avoidance of repetition of adjacent (partially) identical morphs in languages such as Greek, Hungarian, Tswana, Spanish, and Swedish. In section 2.2, I demonstrate the mechanisms of the two approaches with the data from languages such as Mandarin Chinese. I then discuss the problems in several morphologically restricted OT approaches (e.g., de Lacy 1999, Bonet 2004) in section 2.3. In section 2.4, I show how a realization OT approach shows its advantage over a morphologically restricted OT approach in accounting for the relevant data in the languages named above. In section 2.5, I discuss problems caused by voicing assimilation for realization approaches and their possible solutions. I conclude in section 2.6.

2.2 Two types of Optimality-Theoretic approaches

In this section, I discuss the mechanisms of a morphologically restricted OT approach (e.g., de Lacy 1999) and a realization OT approach (e.g., Yip 1998) and demonstrate how they account for cases of (avoidance of) repetition of identical morphs with data from languages such as Mandarin Chinese.

2.2.1 Mandarin *le*

Mandarin Chinese has two types of *le* which are homophonous (Chao 1968, Radford 1977, Li and Thompson 1981, Menn and MacWhinney 1984, Yip 1998, among many others). One is a perfective (PFV) aspectual suffix³:

(4) a. *Ta shui le san ge zhongtou.* (Li and Thompson 1981: 186)
 3sg sleep PFV three CL(ASSIFIER) hour
 ‘S/He slept for three hours.’

b. *Wo zai nali zhu le liang ge yue.* (Li and Thompson 1981: 186)
 I at there live PFV two CL month
 ‘I lived there for two months.’

The other type of *le* is a sentence-final particle indicating a Currently Relevant State (CRS) in Li and Thompson’s terms:⁴

(5) Currently Relevant State

a. *Nei tian ta chu-qu mai dongxi le.* (Li and Thompson 1981: 240)
 that day she exit-go buy thing CRS
 ‘That day she went out shopping.’

b. *Xia-ge yue wo jiu zai Riben le.* (Li and Thompson 1981: 241)
 next-CL month I then at Japan CRS
 ‘Next month I will be in Japan.’

Additionally, the perfective aspectual suffix *-le* can co-occur with the sentence-final particle *le*:

(6) a. *Wo li le fa le.* (Li and Thompson 1981: 201)
 I cut PFV hair CRS
 ‘I (have) had a haircut.’

b. *Wo he le san bei kafei le.* (Li and Thompson 1981: 244)
 I drink PFV three CL coffee CRS
 ‘I have drunk three cups of coffee.’

It is well-known that when the PFV *-le* and the CRS *le* are expected to be adjacent, morphological haplology occurs, i.e., only one *le* shows up instead of the expected **le le* combination.⁵ Each sentence in (7) can have three meanings, i.e., a PFV reading, a CRS reading, or a reading of both PFV and CRS.

(7) *Both-Adjacent (Li and Thompson 1981, cited in Yip 1998: 227)

a. *Huo mie le (*le).*
 fire go out PFV/CRS
 ‘The fire went out (PFV reading) (yesterday).’
 ‘The fire has gone out (CRS reading) (already).’
 ‘The fire went out, and that’s what I’m telling you (PFV and CRS).’

b. *Bing dou hua le (*le).*
 ice all melt PFV/CRS
 ‘The ice all melted.’

2.2.2 A morphologically restricted OT account of Mandarin *le* haplology

De Lacy 1999 provides a thorough account of various types of morphological haplology under a morphologically restricted OT model (McCarthy and Prince 1993b, McCarthy and Prince 1995) and argues that they all can be encoded in constraint-rankings. I demonstrate how de Lacy’s approach accounts for the haplology of Mandarin *le*.

De Lacy assumes that an input to the grammar includes two (partially) identical morphs and proposes the following constraints to account for the haplology of fully identical forms.

(8) a. MAX-IO: Every segment of the input has a correspondent in the output. (No phonological deletion) (McCarthy and Prince 1995)

b. *STRUCTURE: Any overt phonological material is disallowed. (De Lacy 1999)

c. UNIFORMITY: Multiple input elements should not be mapped to a single output correspondent. (No coalescence) (McCarthy and Prince 1995)

De Lacy shows that the ranking schema in (9) accounts for the haplology of fully identical forms.

(9) MAX-IO >> *STRUCTURE >> UNIFORMITY

The ranking schema in (9) can account for the haplology of Mandarin adjacent *le*'s. Consider the tableau in (10). Assume the input to the grammar is comprised of two *le*'s (i.e., $l_1e_2 l_3e_4$). (Since within the morphologically restricted OT model, morphological information such as “affix”, “root”, and “stem” is considered to constitute enough input and output information, we do not need to know which *le* realizes which morphosyntactic feature value as long as we know the two *le*'s are morphs.) Output candidate (a) is ruled out by the highest-ranked constraint MAX-IO in the tableau because l_3e_4 is deleted in the output. Candidate (b) is ruled out because it contains four segments and therefore violates *STRUCTURE four times. Compared to Candidate (b), Candidate (c) only violates *STRUCTURE twice and therefore wins, even if it violates the lowest-ranked constraint UNIFORMITY because it coalesces two input morphs. (The traditional OT symbol \Rightarrow indicates a winning output.)

(10) Haplology of Mandarin *le*

Input: $l_1e_2 l_3e_4$	MAX-IO	*STRUCTURE	UNIFORMITY
a. l_1e_2	*!*	**	
b. $l_1e_2 l_3e_4$		****!*	
\Rightarrow c. $l_{1,3}e_{2,4}$		**	**

2.2.3 A realization OT account

By contrast, Yip 1998 takes a realization OT approach to morphological haplology. Under her approach, inputs including inflectional affixes, clitics, particles (and function words perhaps) are phonologically unrealized, i.e., we find their corresponding morphosyntactic information but not their corresponding phonological information in the input. It is the function GEN that generates all sorts of phonological material to realize the input morphosyntactic information. Additionally, the phonological shape of each inflectional affix, clitic, particle (and function word perhaps) is determined by realization constraints in the grammar.

To account for the haplology of Mandarin *le*, Yip proposes four relevant constraints:

(11) Constraints which account for the haplology of Mandarin *le* (Yip 1998: 228)

- a. PERF: The Perfect verb must end in *le*.
 - b. CRS: Currently Relevant State utterances must end in *le*.
 - c. OCP (*le*): OCP (affix), where affix = *le*.
 - d. MORPHDIS: Distinct instances of morphemes have distinct contents, tokenwise.
- (McCarthy and Prince 1995: 310)

Constraint (11a) says that if an input contains a verb in perfective aspect, then the perfective aspect is realized by the suffix *-le* in the output, assuming that no change of morphosyntactic feature value is made from the input to the output. Constraint (11b) says that if an input contains a CRS discourse feature, then the feature is realized by

the particle *le* in the output. Constraint (11c) is a generalized OCP constraint banning two adjacent *le*-morphs. Constraint (11d) bans a segment which “does double duty to fulfill more than one morphological role” (Yip 1998: 222). (In section 2.2.4 I will use a more straightforward constraint *FEATURE FUSION which bans a morph realizing more than one feature value.)

The grammar of haplology of Mandarin *le* and its corresponding tableau (Yip 1998: 229) are presented in (12) and (13), respectively.

(12) PERF, CRS, OCP (*le*) >> MORPHDIS

Assume the input contains the verb *mie* ‘go out, extinguish’ and the morphosyntactic feature value [+perf(ect)] and the discourse feature CRS. Candidate (a) is ruled out by the grammar because it does not end in *le* and therefore violates the higher-ranked constraints PERF and CRS both of which demand the appearance of *le* in this context. Candidate (b) is ruled out because it violates the higher-ranked constraint OCP (*le*) banning two adjacent *le*’s. Candidate (c) wins, although it violates the lower-ranked constraint MORPHDIS because *le* realizes both [+perf] and CRS.⁶

(13) Tableau

Input: <i>mie</i> , [+perf], CRS	PERF	CRS	OCP (<i>le</i>)	MORPHDIS
a. <i>mie</i> , [+perf], CRS	*!	*		
<div style="text-align: center;"> [+perf] CRS b. <i>mie</i> -<i>le</i> <i>le</i> </div>			*!	
<div style="text-align: center;"> [+perf] CRS \ / c. <i>mie</i> <i>le</i> </div>				*

2.2.4 Cases of repetition of identical morphs

Both the morphologically restricted OT approach and the realization OT approach can account for cases of repetition of identical morphs. Menn and MacWhinney (1984: 528) point out that “[t]olerance of repetition [of identical morphs] ... is the ‘regular’ case: each morpheme of an underlying sequence is marked on the surface by a morph, regardless of the phonetic shapes involved.” They give many cases of repetition of (partially) identical morphs and I only cite a few here:

(14) a. Swahili: The 3pl subject inflectional prefix *wa-* can be adjacent to the 3pl object inflectional prefix *wa-*.⁷

b. Choctaw: The active voice suffix /-li/ can be adjacent to the 1sg suffix /-li/. (Nicklas 1972)

c. Turkish: The past tense suffix /-mVʃ/ can be adjacent to the dubitative clitic /mVʃ/. (Radford 1977)

- d. French: The 1pl subject clitic *nous* can be adjacent to the 1pl object clitic *nous*.
- e. Albanian: The comparative clitic *të* can be adjacent to the 2sg clitic *të*.
- f. German: The feminine definite article *die* can be followed by the feminine relative pronoun *die*.⁸
- g. Tiv: Identical direct and indirect object pronouns may occur in sequence, e.g., /i/ + /i/, /u/ + /u/. (Abraham 1940)

Under de Lacy’s approach, the constraint UNIFORMITY outranks *STRUCTURE in these languages in order to explain repetition of identical morphs. Let us take the Choctaw case as an example. Consider the tableau in (15). Assume the input is comprised of two /-li/’s. Candidate (a) is ruled out because it violates the highest-ranked constraint MAX-IO. Candidate (b) is ruled out because it violates UNIFORMITY in that [l_{1,3}i_{2,4}] coalesces two morphs. Candidate (c) wins although it violates *STRUCTURE four times.

(15) Repetition in Choctaw: A PM account

Input: /-l ₁ i ₂ / + /-l ₃ i ₄ /	MAX-IO	UNIFORMITY	*STRUCTURE
a. l ₁ i ₂	*!*		**
b. l _{1,3} i _{2,4}		*!*	**
☞ c. l ₁ i ₂ + l ₃ i ₄			****

By contrast, we can take a realization OT approach to repetition of identical morphs by assuming that the input contains phonologically unrealized morphosyntactic or semantic feature values. I propose the following constraints:

- (16) a. {active}: /-li/: The active voice feature value is realized by the suffix /-li/.
- b. {1sg}: /-li/: First person singular is realized by the suffix /-li/.
- c. OCP (morph): Two morphs with (partially) identical shapes cannot be adjacent. (“Identity” can be defined with respect to morph initial or final consonants, vowels, etc. depending on a particular language.)⁹
- d. *FEATURE FUSION: A morph cannot realize more than one morphosyntactic or semantic feature value.

Constraints (a) and (b) are realization constraints determining the phonological shape of a morpheme. In this case, they demand the appearance of the suffix /-li/ in the output.

Constraint (c) is a universal constraint which is used to account for avoidance of repetition of identical morphs. OCP (morph) is a generalized constraint (see also Golston 1995, Brentari 1998, Yip 1998)¹⁰ of the original OCP (“Obligatory Contour Principle”) which bans consecutive identical autosegments (Leben 1973, 1978;

Goldsmith 1976, 1984; Pulleyblank 1986), segments (Steriade 1982, Prince 1984, Hayes 1986, Schein and Steriade 1986) or syllables (Yip 1993). The criteria based on which two adjacent morphs are considered identical enough to trigger a violation of OCP (morph) seems to be language-particular: (1) Adjacent morphs with identical initial consonants may violate OCP (morph). For example, Spanish bans two adjacent morphs with initial *l*'s (e.g., **le lo*, **les lo*, **le las*, etc.). (2) Adjacent morphs with identical final consonants may violate OCP (morph). For example, in Swedish, the present tense suffix *-er* is haplologized after a stem ending in *r* (e.g., *rör-* 'move', present tense → *rör*, **rörer* (Stemberger 1981)). (3) Adjacent morphs with identical vowels may violate OCP (morph). In Greek, haplology is obligatory when the past tense prefix *e-* precedes the perfect morpheme *e-* (e.g., *e-* + *e-* + *phthi-meen* → *e-phthi-meen* 'declined (pluperfect)', **e-e-phthi-meen*) (Stemberger 1981, Golston 1995). (4) Sometimes we need to specify the morphological context where OCP (morph) applies. For example, in Romanian, haplology occurs to members of functional categories which are adjacent (17a) but not to a functional marker and a lexical stem which are adjacent (17b) (Ortmann and Popescu 2000: 52-53).

(17) a. *prieten-ul* (**al*) *băiat-ul-ui*
 friend-def.masc poss.sg.masc boy-def.masc.-dat.masc
 'the boy's friend'

b. *cumul* *al* *particul-e-lor* (**cumul particulelor*)
 accumulation poss.sg.masc particle-pl.f-dat.pl
 'accumulation of the particles'

Notice that OCP (morph) is similar to the constraints in Yip 1998 such as OCP (affix), OCP (stem), etc. They are all morphophonological *per se*. Plag 1998, however, argues that haplology can be accounted for by *pure* phonological constraints such as OCP (ONSET) banning two adjacent identical onsets. He uses his constraints to account for cases of morphological truncation in English. Consider the examples in (18) in which the underlined part is truncated. All the illicit forms contain two adjacent identical onsets and are therefore ruled out by OCP (ONSET).

(18) a. *feminine* + *-ize* → *feminize* (**femininize*)
 b. *minimum* + *-ize* → *minimize* (**minimumize*)
 c. *metathesis* + *-ize* → *metathesize* (**metathesize*)

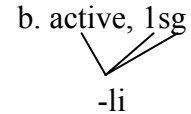
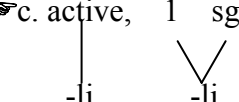
But notice that OCP (ONSET) has to apply in a derived environment and it does not apply to words such as *Titanic*, *dedicate*, etc. Additionally, even in a derived word the application of OCP (ONSET) has to be sensitive to specific affixes (e.g., *irritate* + *-ing* → *irritating* not **irriting*). Therefore, constraints like OCP (ONSET) are still morphologically conditioned constraints.¹¹

Constraint (d) is a markedness constraint banning a form realizing more than one morphosyntactic or semantic feature value. Like the constraint *FEATURE SPLIT which

bans the realization of a morphosyntactic feature value by more than one form, *FEATURE FUSION also favors a model of “one function - one form” (Wurzel 1989).

The grammar which accounts for repetition of identical morphs in Choctaw is shown in (19). According to Yip, no haplology is observed if *FEATURE FUSION (MORPHDIS in Yip’s grammar) outranks OCP (morph). The constraint {1sg}: /-li/ needs to outrank *FEATURE FUSION because /-li/ realizes both 1st person and sg. Assume the input is comprised of unrealized morphosyntactic feature values such as active, 1st person, and sg. Candidate (a) is ruled out because it violates both the realization constraints in that nothing realizes active, 1st person, and sg. Candidate (b) is ruled out because *-li* not only realizes active value but also 1st person and sg, so it violates *FEATURE FUSION twice. (No violation of *FEATURE FUSION is made if there is only one connecting bar between one feature value and one form; one violation is made if there are two bars between two feature values and one form; two violations are made if there are three bars between three feature values and one form.) Candidate (c) violates *FEATURE FUSION once since the following *-li* realizes both 1st person and sg. It also violates OCP (morph) since the two *-li*’s are adjacent. But it still wins.

(19) Repetition in Choctaw: A realization approach

Input: active, 1sg	active: /-li/	{1sg}: /-li/	*FEATURE FUSION	OCP (morph)
a. active, 1sg	*!	*		
b. active, 1sg 			**!	
c. active, 1 sg 			*	*

2.3 Problems for a morphologically restricted OT model

In this section, I discuss some problems for a morphologically restricted OT model to account for (avoidance of) repetition of identical morphs (e.g., de Lacy 1999, Bonet 2004). I show that de Lacy 1999 makes an excessively strong claim that morphological haplology necessarily involves coalescence. Contra de Lacy 1999, I argue that we need OCP-type constraints to account for (avoidance of) repetition of identical morphs. Additionally, I show that in order to account for avoidance of repetition of identical morphs by phonologically unrelated (allo)morph selection, a morphologically restricted OT model has to rely on odd mechanisms.

I first focus on the discussion of de Lacy 1999 which gives a comprehensive OT account of typological cases of (avoidance of) repetition of identical morphs. De Lacy presents an OT account of morphological haplology and makes two claims: (1) “haplology is coalescence”; (2) “there is no generalized OCP.” I show that both the claims are problematic.

As shown in section 2.2.2, de Lacy’s approach to the haplology of Mandarin *le* requires an output *le* which coalesces two input *le*’s. Coalescence, however, is

unconstrained machinery, i.e., any pair of segments can be assumed to coalesce. De Lacy admits that “[coalescence] does not make any prediction about the preservation of features when two affixes haplogogize.”

For example, the Japanese Classical Predicative *-si* [ʃi] haplogogizes with stems ending in both [ʃi] and [ʒi]: e.g., /imizi/ ‘extreme’ + /si/ → [imiʒi], *[imiʒiʃi] (Lawrence 1997). Under de Lacy’s approach, the voiced consonant [ʒ] (in /imizi/) would coalesce with the voiceless consonant [ʃ] (in /si/) while the output would be the voiced consonant [ʒ] (surfaced as [ʒ]).

By contrast, the French noun-forming suffix *-iste* /ist/ haplogogizes with stems ending in /is/ and /iz/: e.g., /analiz/ *analyze* + /ist/ → [analist], *[analizist] (Corbin & Plénat 1992). Under de Lacy’s approach, the voiced consonant [z] (in /analiz/) would coalesce with the voiceless consonant [s] (in /ist/) while the output would be the voiceless consonant [s].¹²

In Swedish, the present tense suffix *-er* haplogogizes after a stem ending in /r/ (e.g., *rör-* ‘move’, present tense *rör*, **rörer*) (Stemberger 1981). Under de Lacy’s approach, *-er* would coalesce with *rör*. It is not clear why this case necessarily involves coalescence.

De Lacy’s coalescence approach often yields odd products. Dressler 1977 remarks that the German suffix *-en* haplogogizes after stems ending in *ern* (e.g., *Eisern* ‘iron’ + *-en* → *eisern*, **eisernen*, **eisen* ‘made of iron’). De Lacy analyzes it as a case of coalescence and accounts for it by ranking *STRUCTURE higher than CONTIGUITY, a constraint “that require[s] retention of underlying adjacency.” Consider the following tableau (from de Lacy 1999). Candidate (a) is ruled out because it contains two more segments than Candidate (b) and therefore causes two extra violations of *STRUCTURE. Candidate (b) wins although it violates the lower-ranked constraint CONTIGUITY because *e*₄ and *n*₅ which are adjacent in the input are not adjacent in the output which has an intervening *r*₂.

(20) German: *Eisern* + *-en*

<i>Eise</i> ₁ <i>r</i> ₂ <i>n</i> ₃ + <i>-e</i> ₄ <i>n</i> ₅	*STRUCTURE	CONTIGUITY
a. <i>eise</i> ₁ <i>r</i> ₂ <i>n</i> ₃ <i>e</i> ₄ <i>n</i> ₅	**!	
☞ b. <i>eise</i> _{1,4} <i>r</i> ₂ <i>n</i> _{3,5}		*

The winning coalesced output *eise*_{1,4}*r*₂*n*_{3,5} is, however, an odd product. What is the status of *en*? It is not a suffix or prefix, which either follows or precedes a stem. It is not an infix, which is surrounded by phonological material on both sides. Neither is it a circumfix, which surrounds other phonological material on both sides. It would be odd to see *en* split if it were a portmanteau morph as described in Anderson (1992) (e.g., in French, *du* ‘of the (= *de* + *le*)’ coalesces more than one morphosyntactic representation; in Breton expressions such as *e dad* ‘his father’, *e zad* ‘her father’, and *tad* ‘father,’ the initial consonant of the noun (/d/ vs. /z/ vs. /t/) can be regarded as functioning both as part of the signal for the possessor and that for the possessed (‘father’)). It would also be odd to see *en* split if it were a submorphemic element like

gl- in English (e.g., *glitter*, *gleam*, *glow*, etc.). Additionally, the sandwiched *r* is not an epenthesized element to satisfy phonotactic constraints and *en* is a German suffix (unlike Semitic languages in which root consonants are interwoven with vowels).

Moreover, in his paper de Lacy repeatedly claims that there is no OCP-like constraint. However, it is not clear whatsoever how his approach based on coalescence and any relevant markedness constraint (except OCP) would account for identity-triggered selection of phonologically unrelated (allo)morphs. For example, in Hungarian the 2nd person singular indefinite present indicative marker *-sz* ([s]) is replaced by *-ol* when it appears after stems ending in sibilants and affricates (Carstairs 1988, 1990). It is not clear which relevant markedness constraint (except OCP) prefers [l] to [s].

Bonet 2004 advocates a morphologically restricted OT approach to cases of phonologically conditioned allomorph selection in which the input can contain more than one allomorph. Under Bonet’s approach, the Hungarian morphs *-sz* and *-ol* would be placed in an input set in which *-sz* has priority over *-ol* in being spelt out. The input set can be formulated as $\{-sz > -ol\}$.¹³ Bonet also proposes “[a] universal constraint called PRIORITY [which] ensures that this preference relation is obeyed” (p.90). The grammar can be formulated as “OCP (morph) >> PRIORITY.” Consider the tableaux in (21). In (21a), Candidate (a) is ruled out by PRIORITY. In (21b), Candidate (b) is ruled by OCP (morph) which bans two adjacent partially identical morphs because both *olvas* and *-sz* have a final strident. Candidate (a) (*olvas-ol*) wins though it violates PRIORITY.

(21) a. Hungarian: *ír-ni* ‘to write’ (*-ni* marks an infinitive)

<i>ír</i> + $\{-sz > -ol\}$	OCP (morph)	PRIORITY
a. <i>ír-ol</i>		*!
☞ b. <i>ír- sz</i>		

b. Hungarian: *olvas-ni* ‘to read’

<i>olvas</i> + $\{-sz > -ol\}$	OCP (morph)	PRIORITY
☞ a. <i>olvas-ol</i>		*
b. <i>olvas-sz</i>	*!	

This approach has several problems. It is odd to put a phonological representation like $\{-sz > -ol\}$ in the input. The set $\{-sz > -ol\}$ basically says that only one input can show up in the output and *-sz* has priority over *-ol*, which violates the tenet of Optimality Theory that there should be no input grammar. Additionally, it is a pure stipulation that in an input set which has more than one (allo)morph, only one can show up in the output. It is not clear how a morphologically restricted OT model rules out candidates like **olvas-ol-sz* without introducing morphosyntactic feature values. Moreover, the nature of the constraint PRIORITY is not clear. It seems to be a faithfulness constraint which spells out the information encoded in the input. However, faithfulness constraints usually operate on representations. By contrast, PRIORITY is a constraint faithful to a relation that anything on the left side of “>” has priority in

being spelt out.

2.4 A realization OT approach to avoidance of repetition of (partially) identical morphs

In this section, I show that the above-mentioned problems caused by coalescence of adjacent partially identical morphs without clear evidence, the stipulation that in an input set which has more than one (allo)morph, only one can show up in the output, and constraints like PRIORITY which spell out the stipulation in the input are readily solved in a realization OT approach to avoidance of repetition of (partially) identical morphs. I demonstrate the advantages of the realization approach with data from several languages such as Greek, Hungarian, Tswana, Spanish, and Swedish.

2.4.1 Greek negatives

Golston 1995 discusses a case of OCP-triggered allomorph selection in Greek. The negative *meé* ‘not’ is “used in irrealis contexts primarily governing optative (22a), subjunctive (22b) and imperative (22c) verb forms” (p.358):

- (22)a. *eè meè dzoó-ieven*
or not live-1OPT
‘or may I not live’
- b. *meè phóo-men*
not say-1P SUBJ
‘shall we not say?’
- c. *meè még-a lég-e*
not big-A:N:P speak-2IMP
‘don’t boast’

The string *meé* can also be a subordinating conjunction (‘lest’) and introduce subordinate clauses like the object clauses used with verbs of fearing (23) (p.358):

- (23) *dé-doi-ka meè...epilathoó-metha t-ées oík-ade hod-óu*
REDUP-fear-1PERF lest lost-1P SUBJ the-G:F homeward road-G:F
‘I fear we may forget the way home’

Interestingly, when the two *meé*’s are expected to be adjacent, the negative *meé* is replaced by its allomorph *ou* (*ouk* before a vowel initial word¹⁴) which generally occurs “in realis contexts primarily governing indicative forms of the verb” (p.358). (24) is a case in which *ou(k)* is used in a realis context.

- (24) *ouk en-no-óo* (Golston 1995: 358)
not in-mind-1IND
‘I don’t recall’

In (25) (Golston 1995: 358), the context is irrealis as indicated by the verb *ée-te* ‘you are’ which appears in the subjunctive mood. As we can see, *ou* ‘not’ which is supposed to be in a realis context shows up after *meé* so that the combination of two adjacent *meé*’s is avoided.

- (25) *dé-di-men* *meé* *ou* *bébai-oi* *ée-te*
 REDUP-fear-1P INDIC lest not steady-M:P be-2P SUBJUNCTIVE
 ‘We fear you are not to be depended on.’

To account for the above data, I propose the realization constraints in (26) apart from OCP (morph) which bans two adjacent identical morphs (in this case). The constraints in (26) indicate a subset relation, i.e., *ou* is treated as a default negative and realizes a subset of the feature values realized by *meé*.

- (26) a. {[+irrealis], negative}: *meé*: The features [+irrealis] and “negative” are realized by *meé*.
 b. Negative: *ou*: The feature “negative” is realized by *ou*.

The ranking of the two realization constraints follows the well-known “elsewhere condition” (Kiparsky 1973, Anderson 1986, among many others) which can also be called “the specificity condition”. The constraint realizing *meé* outranks the one realizing *ou* because the former constraint is morphemically more specific and realizes both [+irrealis] and negative while the latter constraint only realizes negative. Therefore, *meé* has priority over *ou* in being spelt out. The two realization constraints are outranked by OCP (morph)¹⁵ and *FEATURE SPLIT banning the realization of a morphosyntactic or semantic feature value by more than one form.

Consider the tableaux in (27). In (27a), assume the input is comprised of phonologically unrealized features [-irrealis] and “negative”. Candidate (a) is ruled out by the constraint realizing *ou* because there is no *ou* in the output to realize “negative”. Candidate (b) is ruled out in violation of *FEATURE SPLIT. Candidate (c) wins without violating any of the constraints in (27a). Notice that Candidate (c) vacuously satisfies the constraint {[+irrealis], negative}: *meé*.¹⁶ In (27b), assume the input contains unrealized features such as [+irrealis] and “negative” and the form realizing these features will be adjacent to *meé* ‘lest’. Candidate (a) is ruled out by OCP (morph). Candidate (b) is ruled out in violation of *FEATURE SPLIT because ‘negative’ is realized by both *meé* and *ou*. Candidate (c) wins despite its violation of the constraint realizing the negative *meé*.

(27) Tableaux: Greek negatives

a. *ou*

[-irrealis], negative	OCP (morph)	*FEATURE SPLIT	[+irrealis], negative: <i>meé</i>	negative: <i>ou</i>
a. [-irrealis], negative <i>meé</i>				*!
b. [-irrealis], negative <i>meé ou</i>		*!		
☞ c. [-irrealis], negative <i>ou</i>				

b. *meé ou*

<i>meé</i> ‘lest’, [+irrealis], negative	OCP (morph)	*FEATURE SPLIT	[+irrealis], negative: <i>meé</i>	negative: <i>ou</i>
a. ‘lest’, [+irrealis], negative <i>meé meé</i>	*!			*
b. ‘lest’, [+irrealis], negative <i>meé ou meé</i>		*!		
☞ c. ‘lest’, [+irrealis], negative <i>meé ou</i>			*	

The avoidance of two adjacent *meé*'s in Greek is a case of “the emergence of the unmarked” (McCarthy 1994). In the context of [+irrealis], when a negative is not adjacent to *meé* ‘lest’, the effect of OCP (morph) is not observed and therefore the more specific negative morpheme *meé* is realized. However, when a negative is adjacent to *meé* ‘lest’, OCP (morph) forces the default negative *ou* to emerge even in the context of [+irrealis].

The proposed realization OT approach not only captures this observation but also follows a particularly important principle (i.e., the specificity condition) that autonomously arranges constraint rankings. By contrast, a morphologically restricted OT approach may simply stipulate the relation in the input that *meé* has priority over *ou* in being spelt out and that only one morph, either *meé* or *ou*, can be spelt out.

2.4.2 Hungarian second person singular indefinite

Another case of “the emergence of the unmarked” triggered by OCP is found in Hungarian. Hungarian basically has two types of conjugations, both definite and indefinite. “The definite conjugation is used if the sentence contains a definite direct

object. The indefinite conjugation is used at all other times” (Rounds 2001: 23). The 2nd person singular indefinite present indicative marker is the suffix *-sz* ([s]) (e.g., *ír-sz* ‘(you sg) write (something [-definite])’). Interestingly, when *-sz*¹⁷ is adjacent to a verbal stem ending in a strident (*s* [ʃ], *sz* [s], *z* [z], *dz* [ʒ]), it is replaced by *-Vl* (subject to vowel harmony) (e.g. *olvas-ol* ‘(you sg) read (something [-definite])’, **olvas-(o)sz*) (Vago 1980, Carstairs 1988, 1990, Rounds 2001, among others). Additionally, *-Vl* is a more general and default marker for 2nd person sg indefinite compared to *-sz* because *-Vl* not only marks present indicative but also past, conditional, and subjunctive.¹⁸

(28) Hungarian 2nd person sg indefinite conjugation (adapted from Vago 1980: 50)

	/va:r/ ‘wait for’	/hoz/ ‘bring’	/ad/ ‘give’
Present	<i>vársz</i>	<i>hozol</i>	<i>adsz</i>
Past	<i>vártál</i>	<i>hoztál</i>	<i>adtál</i>
Conditional	<i>várnál</i>	<i>hoznál</i>	<i>adnál</i>
Subjunctive	<i>várjál</i>	<i>hozzál</i>	<i>adjál</i>

To account for the *-sz/-Vl* alternation in (28), I propose the relevant constraints in (29). The ranking of (29a) and (29b) follows the specificity condition with (29a) outranking (29b).¹⁹ Constraint (29d) requires *-sz* to be adjacent to a verbal stem (without any intervening material). It is attested by the examples like *vársz* and *adsz* in (28) (**városz*, **adosz*).

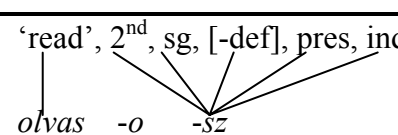
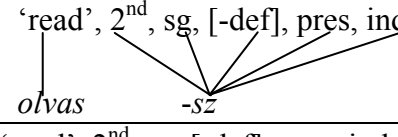
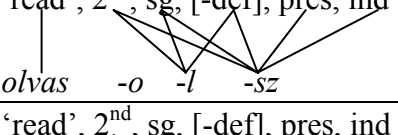
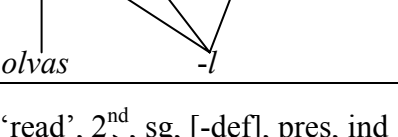
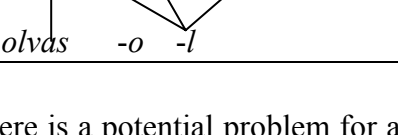
- (29) a. {2nd person, sg, indefinite, present, indicative}: *-sz*: The feature value set {2nd person, sg, indefinite, present, indicative} is realized by *-sz*. (*-sz*)
 b. {2nd person, sg, indefinite}: *-Vl*: The feature value set {2nd person, sg, indefinite} is realized by *-Vl*. (*-Vl*)
 c. OCP (strident): No adjacent stridents are allowed.
 d. ALIGN (Verbal stem, Right, {2nd person, sg, indef, pres, ind}, Left): The right edge of a verbal stem coincides with the left edge of the marker of {2nd person, sg, indef, pres, ind} (i.e., *-sz*). (ALIGN (V, R, *sz*, L))
 e. *FEATURE SPLIT: A morphosyntactic or semantic feature value cannot be realized by more than one form. (*FS)

There are several ways to formulate a constraint like (29b). {2nd person, sg, indefinite} could be realized by *-l* ({2nd person, sg, indefinite}: *-l* (*-l*)) with a vowel inserted by other constraints (*o* [o] after a stem, *a* [ɔ] after an inflectional suffix, and *á* [a:] between an inflectional suffix and *l*) (Vago 1980, Siptár and Törkenczy 2000).

Consider the tableau in (30). Assume the input is comprised of the verbal stem *olvas-* and the phonologically unrealized morphosyntactic feature value set {2nd person, sg, indefinite, present, indicative}. In addition to the above-mentioned constraints, I propose a markedness constraint *[sl](-CODA) given that Hungarian does not allow a coda with a strident followed by [l] (Siptár and Törkenczy 2000: 106). Assume the grammar is *[sl], OCP (strident), *FEATURE SPLIT, ALIGN (V, R, *sz*, L) >>

-sz >> *-l*. Candidate (a) is ruled out because it is not adjacent to the verbal stem *olvas-* and therefore violates ALIGN (V, R, *sz*, L). Candidate (e) wins though it violates the constraint realizing *-sz*. Notice that in this case ALIGN (V, R, *sz*, L) can rule out Candidate (c) without *FEATURE SPLIT. We will further discuss the distinction between alignment constraints and *FEATURE SPLIT in Chapter 4.

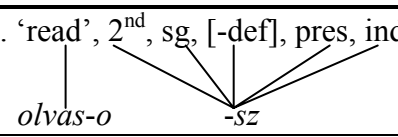
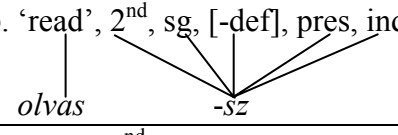
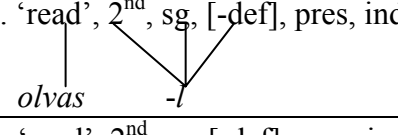
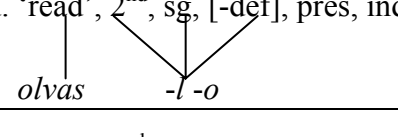
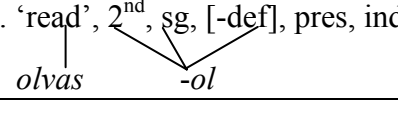
(30) The *-sz/-l* alternation

<i>olvas-</i> ‘read’, 2 nd , sg, [-def], pres, ind	*[s]	OCP	*FS	ALIGN (V, R, <i>sz</i> , L)	<i>-sz</i>	<i>-l</i>
a. ‘read’, 2 nd , sg, [-def], pres, ind 				*!		*
b. ‘read’, 2 nd , sg, [-def], pres, ind 		*!				*
c. ‘read’, 2 nd , sg, [-def], pres, ind 			**!*	**		
d. ‘read’, 2 nd , sg, [-def], pres, ind 	*!					*
e. ‘read’, 2 nd , sg, [-def], pres, ind 					*	

There is a potential problem for a formulation like {2nd person, sg, indef}: *-l*. If a vowel can be inserted between a stem and *-l*, why can't it be inserted after *-l*? The grammar in (30) will also choose the incorrect candidate **olvas-l-o*. The illicit form **olvas-l-o* does not violate any phonotactic constraint because Hungarian allows intervocalic consonant clusters with a strident followed by [l] (Siptár and Törkenczy 2000: 129-130). We may use an alignment constraint to make *-l* appear at the rightmost edge so that **olvas-l-o* is ruled out.

We can also assume that {2nd person, sg, indefinite} is realized by a morph which already contains a vowel in addition to [l]. We can assume, for example, *-ol* is the marker of {2nd person, sg, indefinite} (Carstairs 1988, 1990) and posit the realization constraint {2nd person, sg, indefinite}: *-ol*. The grammar in (31) chooses the desired output candidate *olvas-ol*. (*[s] becomes irrelevant.) Candidates (c) and (d) are ruled out because they violate both the constraints realizing *-sz* and *-ol*.

(31) The *-sz/-ol* alternation

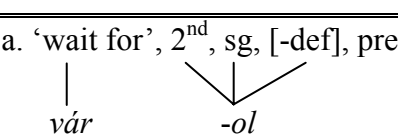
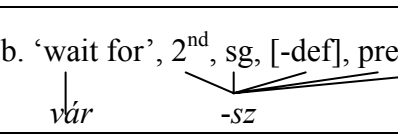
<i>olvas-</i> ‘read’, 2 nd , sg, [-def], pres, ind	OCP	ALIGN (V, R, <i>sz</i> , L)	<i>-sz</i>	<i>-ol</i>	DEP (V)
a. ‘read’, 2 nd , sg, [-def], pres, ind 		*!		*	*
b. ‘read’, 2 nd , sg, [-def], pres, ind 	*!			*	
c. ‘read’, 2 nd , sg, [-def], pres, ind 			*	*!	
d. ‘read’, 2 nd , sg, [-def], pres, ind 			*	*!	*
e. ‘read’, 2 nd , sg, [-def], pres, ind 			*		*

A question arises. Why should {2nd person, sg, indefinite} be realized by *-ol*? There are several morphophonologically related allomorphs of *-Vl*. The suffix *-ol* occurs after back vowels; *-el* after front unrounded vowels; *-öl* after front rounded vowels (Rounds 2001: 26). Outside {2nd person, sg, indefinite, present, indicative}, *-ál* occurs after an inflectional suffix. We need a separate grammar to derive these allomorphs if we assume {2nd person, sg, indefinite} is realized by *-ol*.

We can also assume that {2nd person, sg, indefinite} is realized by *-Vl* (i.e., Constraint (29b)), a suffix which contains a vowel whose phonological features are underspecified. Thus, any output vowel plus [l] will satisfy this constraint and the quality of the output vowel is determined by other constraints conditioning vowel harmony or lengthening. The grammar in (31) with *-ol* replaced by *-Vl* will also choose the correct output *olvas-ol* and rule out **olvas-(o)-sz*.

This grammar will correctly choose a candidate with *-sz* and rule out the one with *-Vl* when the input is comprised of {2nd person, sg, indefinite, present, indicative} and a verbal stem which does not end in a strident. See the following tableau.

(32) *vár-* /va:r/ ‘wait for’, present

<i>vár-</i> ‘wait for’, 2 nd , sg, [-def], pres, ind	OCP	ALIGN (V, R, <i>sz</i> , L)	<i>-sz</i>	<i>-Vl</i>	DEP (V)
a. ‘wait for’, 2 nd , sg, [-def], pres, ind 			*!		*
b. ‘wait for’, 2 nd , sg, [-def], pres, ind 				*	

Additionally, the grammar in (32) and (33) correctly chooses a candidate with *-Vl* and rules out the one with *-sz* when the input comprises {non-present, 2, sg, indefinite}. Consider the tableau in (33). The past tense is realized by the suffix *-t*. (The constraint [+past]: *-t* is not in (33) and we will discuss the linear order of morphs with respect to their morphosyntactic feature values in Chapter 3.) Candidate (b) wins and vacuously satisfies the constraint realizing *-sz*. Notice that *-asz* occurs after two consonants or a long vowel plus [t] (Rounds 2001: 26). The *-sz/-asz* alternation, however, does not affect the result that *-Vl* wins because both *-sz* and *-asz* are ruled out by the constraint realizing *-Vl*. Candidate (a) does not violate ALIGN (V, R, *sz*, L) because *-sz* in Candidate (a) does not realize {2, sg, [-def], present, indicative}.

(33) *vár-* /va:r/ ‘wait for’, past

<i>vár-</i> ‘wait for’, 2, sg, [-def], past	OCP	ALIGN (V, R, <i>sz</i> , L)	<i>-sz</i>	<i>-Vl</i>	DEP (V)
a. ‘wait for’, past, 2, sg, [-def] <div style="text-align: center;"> <i>vár</i> <i>-t</i> <i>-sz</i> </div>				*!	
b. ‘wait for’, past, 2, sg, [-def] <div style="text-align: center;"> <i>vár</i> <i>-t</i> <i>-ál</i> </div>					*

It is not our major concern whether to realize {2, sg, indefinite} by *-l*, *-ol* or *-Vl*. Crucially, the realization approach in question readily accounts for the alternation of *-sz/-Vl* triggered by OCP. It follows the specificity condition which determines the ranking of *-sz* and *-Vl*. It provides a consistent grammar which chooses the right output in various contexts. By contrast, a morphologically restricted OT account has to stipulate in the input that either *-sz* or *-Vl* can be spelt out and *-sz* has priority in the context of {2, sg, indefinite, present, indicative} (e.g., {-*sz* > *-l*}), while it has to posit a different input (e.g., *-l*) in the context of {2, sg, indefinite, non-present}.

2.4.3 Tswana predicative concords

The OCP-triggered morph selection is also found in Tswana, a widely dispersed Bantu language in Southern Africa. Tswana has “two types of predicative concords, subjectival and objectival” (Cole 1955: 229). The normal subjectival concords (SCs) which are used in the principal tenses of the primary moods are shown as follows.

(34) Tswana normal subjectival concords (Cole 1955: 230)

		Singular	Plural
1 st pers.		<i>ke-, n-</i>	<i>re-</i>
2 nd pers.		<i>ò-</i>	<i>lò- [lè]</i>
3 rd pers. cl.	1	<i>ó-</i>	<i>ba-</i>
	2	<i>ó-</i>	<i>e-</i>
	3	<i>lé-</i>	<i>a-</i>
	4	<i>se-</i>	<i>di-</i>
	5	<i>e-</i>	<i>di-</i>
	6	<i>ló-</i>	<i>di-</i>
	7	<i>bo-</i>	<i>a-</i>
	8, 9	<i>go-</i>	

The normal objectival concords (OCs) are shown as follows.

(35) Tswana normal objectival concords (Cole 1955: 231)

		Singular	Plural
1 st pers.		<i>N-</i>	<i>re-</i>
2 nd pers.		<i>go-</i>	<i>lo- [le-]</i>
3 rd pers. cl.	1	<i>mo-</i>	<i>ba-</i>
	2	<i>o-</i>	<i>e-</i>
	3	<i>le-</i>	<i>a-</i>
	4	<i>se-</i>	<i>di-</i>
	5	<i>e-</i>	<i>di-</i>
	6	<i>lo-</i>	<i>di-</i>
	7	<i>bo-</i>	<i>a-</i>
	8, 9	<i>go-</i>	

As pointed out by Cole, “the OCs are identical in form with the [normal] SCs except in the 1st and 2nd persons singular and in class 1 singular” (p.231). The normal 1st person sg SC is generally marked by the prefix *ke-*. In (36), the prefix *a-* is a principal formative used in various indicative forms (PrInd), i.e., *a-* is a marker of the indicative mood in main clauses.

(36) *ke-a-rêka*²⁰ (Cole 1955: 244)
 1sg SC-PrInd-buy
 ‘I buy, I am buying’

The 1sg OC is marked by the syllabic nasal prefix *N-* which agrees with its following consonant in place.

(37) a. *-bôna* ‘see’ → *Ó-a-m-pôna* (Cole 1955: 41)
 3sg SC-PrInd-1sg OC-see
 ‘He sees me.’

- b. *-dirêla* ‘do’ → *go-n-tirêla* (Cole 1955: 41)
 Infin-1sg OC-do
 ‘to do for me’

Interestingly, when the 1sg SC *ke-* is expected to precede the conditional mood (CM) affix *ka-* or a stem beginning with *k-* (e.g., *-kile* ‘once’ (a deficient or auxiliary verbal stem)), it is usually replaced by the nasal prefix (velar in agreement with *k-*) which is also a 1sg OC (Cole 1995: 230, Menn and MacWhinney 1984). See (38).

- (38) a. *N-ka-rêka* (Cole 1955: 230)
 1sg SC-CM-buy
 ‘I would/can buy.’

- b. *N-kilê ka-rêka* (Cole 1955: 297)
 1sg SC-PERF 1sg SC-buy
 ‘I once bought.’

To account for the *ke-/N-* alternation, as we did in Greek and Hungarian, we can assume that *ke-* is a marker of {1, sg, SC} while *N-* potentially marks both {1, sg, SC} and {1, sg, OC}, and is therefore a marker of {1, sg}.

Additional evidence that *N-* is a default marker of {1, sg} includes that *N-* appears in the 1sg absolute pronoun *nna*. According to Cole, Tswana pronouns “are ‘concord’ converted into complete words” (p.127) and “[t]hey may stand alone, as subject or object in a sentence” (p.128). See (39) for a table of Tswana absolute pronouns. “[T]he absolute pronouns consist of a radical portion which usually contains the vowel *ô*, together with a suffixal element *-na* or *-nê* [which] is not a stem, but a stabilizer, whose sole function is to provide a second syllable and thereby to avoid the monosyllabic words which would otherwise result, and to which Tswana, like other Bantu languages, is antipathetic to a greater or less extent” (p.128).

- (39) Tswana absolute pronouns (Cole 1955: 128)

	Singular	Plural
1 st pers.	<i>nna</i>	<i>rona</i> [<i>tšhona</i>]
2 nd pers.	<i>wêna</i>	<i>lona</i> [<i>nyena, lena</i>]
3 rd pers. cl. 1	<i>ênê</i> [<i>yênê</i>]	<i>bônê</i>
2	<i>ônê</i> [<i>wônê</i>]	<i>yônê</i>
3	<i>lônê</i> [<i>jônê</i>]	<i>ônê</i>
4	<i>šônê</i>	<i>tšônê</i>
5	<i>yônê</i>	<i>tšônê</i>
6	<i>lônê</i>	<i>tšônê</i>
7	<i>jônê</i> [<i>bônê, bjônê</i>]	<i>ônê</i>
8, 9	<i>gônê</i>	

Additionally, *n* is part of the marker of 1sg quantitative concord. See (40) for a table of quantitative concords. These concords can attach to the “exclusive” stem *-si* ‘only, alone’. Except 1st and 2nd persons singular and class 1 singular, they can also attach to the “inclusive” stem *-tlhê* ‘the whole, all’. “The quantitative concords have a characteristic vowel *ô* except in the 2nd and 3rd persons singular where it is *ê*” (Cole 1955: 155).

(40) Quantitative concords (Cole 1955: 154)

		Singular	Plural
1 st pers.		<i>nô-</i>	<i>rô-</i>
2 nd pers.		<i>wê-</i>	<i>lô-</i>
3 rd pers. cl.	1	<i>ê-</i> [yê-]	<i>bô-</i>
	2	<i>ô-</i> [wô-]	<i>yô-</i>
	3	<i>lô-</i> [jô-]	<i>ô-</i>
	4	<i>šô-</i>	<i>tšô-</i>
	5	<i>yô-</i>	<i>tšô-</i>
	6	<i>lô-</i>	<i>tšo-</i>
	7	<i>jô-</i> [bô-, bjô-]	<i>ô-</i>
	8, 9	<i>gô-</i>	

To account for the *ke-/N-* alternation, I propose the following constraints.

- (41) a. Conditional Mood: *ka-*: Conditional Mood is marked by *ka-*. (CM: *ka-*)
 b. Principle Indicative: *a-*: The indicative mood in main clauses is marked by *a-*. (PrInd: *a-*)
 c. OCP (morph): Two morphs with identical initial consonants should not be adjacent.
 d. {1, sg, Subject Concord}: *ke-*: {1, sg, SC} is marked by *ke-*.
 e. {1, sg}: *N-*: {1, sg} is marked by a syllabic nasal.

The grammar is CM: *ka-* >> PrInd: *a-*, OCP (morph) >> {1, sg, SC}: *ke-* >> {1, sg}: *N-*. The ranking of {1, sg, SC}: *ke-* and {1, sg}: *N-* follows the specificity condition.

Consider the tableau in (42). Let us put aside the issue of what constraints determine the linear order of the prefixes in question. In (42a), Candidate (a) is ruled out because it violates the higher-ranked constraint realizing *ke-*. In (42b), Candidate (a) is ruled out because the marker of the Conditional Mood (*ka-*) is not spelt out. Candidate (b) in (42b) and Candidate (a) in (42c) lose in violation of OCP (morph).

(42) a. *ke-a-rêka* ‘I buy, I am buying’

1, sg, SC, Ind (main clause), <i>rêka</i> ‘buy’	CM: <i>ka-</i>	PrInd: <i>a-</i>	OCP (morph)	1sg SC: <i>ke-</i>	1sg: <i>N-</i>
a. 1 sg SC, Ind, ‘buy’ 				*!	
☞ b. 1 sg SC, Ind, ‘buy’ 					*

b. *N-ka-rêka* ‘I would/can buy.’

1, sg, SC, Conditional Mood, <i>rêka</i> ‘buy’	CM: <i>ka-</i>	PrInd: <i>a-</i>	OCP (morph)	1sg SC: <i>ke-</i>	1sg: <i>N-</i>
a. 1 sg SC, CM, ‘buy’ 	*!				*
b. 1 sg SC, CM, ‘buy’ 			*!		*
☞ c. 1 sg SC, CM, ‘buy’ 				*	

c. *N-kilê ka-rêka* ‘I once bought.’²¹

1, sg, SC, Subjunctive Mood, <i>kilê</i> ‘once’	CM: <i>ka-</i>	PrInd: <i>a-</i>	OCP (morph)	1sg SC: <i>ke-</i>	1sg: <i>N-</i>
a. 1 sg SC, SM, ‘once’ 			*!		*
☞ b. 1 sg SC, SM, ‘once’ 				*	

The ranking OCP (morph) >> {1, sg, SC}: *ke-* >> {1, sg}: *N-* shows another case of the emergence of the unmarked. In the context of 1sg subject concord, *N-* which potentially marks both 1sg subject and object concords generally does not appear because of the higher-ranked constraint realizing *ke-*. However, in cases where a violation of OCP is triggered, the lower-ranked *N-* emerges to avoid a combination of two adjacent morphs with initial *k*'s. By contrast, a morphologically restricted OT account of the *ke-/N-* alternation will have to stipulate different input representations in different morphosyntactic contexts.

Notice that concord markers such as *ba-* {3, plural} can be adjacent to a morph with an identical initial consonant. See (43).

- (43) *botsa* ‘ask’ → *go-ba-botsa*²² (Cole 1955: 239)
 Infinitive-3PL-ask
 ‘to ask them (i.e., *batho* ‘people’)

To account for cases like (43), we can say that OCP (morph) is a morpheme-specific markedness constraint and only bans two adjacent morphs with initial *k*'s. We can also assume that the constraint realizing *ba-* outranks OCP (morph) which bans two adjacent morphs with any identical initial consonants: {3, plural}: *ba-* >> OCP (morph). Notice that {3, plural}: *ba-* can outrank {1, sg, SC}: *ke-* since there is no subset relation between the two realization constraints. Consider the tableau in (44). Despite its violation of OCP (morph), Candidate (b) wins because *ba-* needs to be spelt out.

- (44) *go-ba-botsa* ‘to ask them’

Infinitive, 3pl Object Concord, <i>botsa</i> ‘ask’	{3, plural}: <i>ba-</i>	OCP (morph)
a. Infinitive, 3pl OC, ‘ask’ <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> go- </div> <div style="text-align: center;"> botsa </div> </div>	*!	
☞ b. Infinitive, 3pl OC, ‘ask’ <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> go- </div> <div style="text-align: center;"> ba- </div> <div style="text-align: center;"> botsa </div> </div>		*

2.4.4 Spanish clitics

The same analysis which we made of the Greek, Hungarian, and Tswana data applies to Spanish clitics.²³ In Standard Spanish, the clitic *lo* marks 3rd person singular non-reflexive masculine direct object. In (45), *lo* can refer to the direct objects *el libro* ‘the book’ (45b) and *a Pedro* ‘to Peter’ (45c) in Standard Spanish.²⁴

- (45) a. *Juan lo vió.*
 John it see-3SG.PAST
 ‘John saw it/him.’
- b. *Juan vió el libro.*
 John see-3SG.PAST the book
 ‘John saw the book.’
- c. *Juan vió a Pedro.*
 John see-3SG.PAST to Peter
 ‘John saw peter.’

In Standard Spanish, the clitic *le* marks 3rd person sg non-reflexive indirect

objects, either masculine or feminine.²⁵ In (46), *le* can refer to both the indirect objects *a Pedro* ‘to Peter’ and *a María* ‘to Mary’.

- (46) a. *Juan le dió el libro.*
 John him/her gave the book
 ‘John gave the book to him/her.’
- b. *Juan dió el libro a Pedro / a María.*
 John gave the book to Peter / to Mary
 ‘John gave the book to Peter/Mary.’

Additionally, the clitic *la* is a marker of 3rd sg non-reflexive feminine direct object, either human or non-human. In (47), *la* can refer to both the direct objects *la mesa* ‘the table (feminine)’ and *a Susana* ‘Susana’.

- (47) a. *Juan la vió.*
 John it/her see-3SG.PAST
 ‘John saw it/her.’
- b. *Juan vió la mesa.*
 John see-3SG.PAST the table
 ‘John saw the table.’
- c. *Juan vió a Susana.*
 John see-3SG.PAST to Susana
 ‘John saw Susana.’

In Spanish, *se* can be used as a 3rd person reflexive clitic, either singular (48a) or plural (48b), or an impersonal pronoun (48c).

- (48) a. *Juan_i se_{i(*j)} vió.*
 John himself see-3SG.PAST
 ‘John saw himself.’
- b. *Juan se lava las manos.*
 John REFLEXIVE wash-3SG.PRESENT.INDICATIVE the hands
 ‘John washes his hands.’
- c. *Se venden flores en el pueblo.*
 IMPERSONAL PRONOUN sell-3PL.PRESENT flowers in the town
 ‘People sell flowers in the town.’

One of the most interesting and well-known phenomena in Spanish is the spurious *se* rule (Perlmutter 1971, Radford 1977, Bonet 1991, 1995, Grimshaw 1997b,

2001, among others). When two clitics with initial *l*'s are adjacent, one of them, which is usually *le* or its plural form *les*, is replaced by *se*. See (49).

(49) a. *Juan se (*le) lo dió.*
 John him/her it/him give-3SG.PAST
 'John gave it/him to him/her.'

b. *Juan se (*le) la dió.*
 John him/her it/her give-3SG.PAST
 'John gave it/her to him/her.'

To account for the spurious *se* rule in OT, apart from *FEATURE SPLIT (*FS), I propose the markedness constraint OCP *lV(C)*:

(50) OCP *lV(C)*: Two clitics with initial *l*'s should not be adjacent.²⁶

OCP *lV(C)* applies to two adjacent clitics. It is a morphophonological constraint in nature rather than a pure phonological one (cf., Plag 1998). Monomorphemic words with adjacent *l*-onsets exist in Spanish (51).

(51) a. *lelo* 'stupid' (colloquial) b. *Lola* (a girl's name) c. *Lalo* (a last name)

An enclitic with an initial *l* can also be adjacent to an *l*-onset in a verbal stem (52b).

(52) a. *Muele el café!*
 Grind-IMPERATIVE the coffee
 'Grind the coffee!'

b. *Muélelo.*
 'Grind it!'

Moreover, I propose the following relevant realization constraints:

- (53) a. {3, sg, [-ref(lexive)], acc, masculine, object}: *lo*: The feature value set {3, sg, [-ref], acc, masc, object} is realized by *lo*. (LO)
 b. {3, sg, [-ref], acc, feminine, object}: *la*: The feature value set {3, sg, [-ref], acc, feminine, object} is realized by *la*. (LA)
 c. {3, sg, [-ref], object}: *le*: The feature value set {3, sg, [-ref], object} is realized by *le*. (LE)
 d. Default: *se*: Every morphosyntactic feature value should be realized by *se*.²⁷
 (SE)

The ranking of the above realization constraints follows the specificity condition, with the higher-ranked constraint realizing a more specific feature value set (54).

(54) LO, LA >> LE >> SE

Consider the tableaux in (55). In (55a), Candidates (a), (b), and (d) are ruled out because they all violate the constraint realizing *le*. The winning Candidate (c) vacuously satisfies the constraints realizing *lo* and *la*, and violates the lowest-ranked constraint realizing *se*. Candidate (e) is ruled out in violation of *FEATURE SPLIT because both *se* and *le* appear to realize one feature value set although *se* and *le* can co-occur when they realize two distinct feature value sets (56). In (55b), Candidate (a) is ruled out by OCP IV(C) because it has two adjacent clitics with initial *l*'s. Candidate (b) loses in violation of the higher-ranked constraint realizing *lo*.

(55) Spanish clitics

a. *Juan le dió el libro.* ‘John gave the book to **him/her**.’

{3 sg [-ref] dat object}	OCP IV(C)	*FS	LO	LA	LE	SE
a. {3 sg [-ref] dat object}: <i>lo</i>					*!	*
b. {3 sg [-ref] dat object}: <i>la</i>					*!	*
☞ c. {3 sg [-ref] dat object}: <i>le</i>						*
d. {3 sg [-ref] dat object}: <i>se</i>					*!	
e. {3 sg [-ref] dat object}		*!***				

b. *Juan se lo dió.* ‘John gave it (**masc**) to **him/her**.’

{3 sg [-ref] dat object} & {3 sg [-ref] acc masc object}	OCP IV(C)	*FS	LO	LA	LE	SE
a. {3 sg [-ref] dat object}: <i>le</i> {3 sg [-ref] acc masc object}: <i>lo</i>	*!				*	*
b. {3 sg [-ref] dat object}: <i>se</i> {3 sg [-ref] acc masc object}: <i>le</i>			*!		*	*
☞ c. {3 sg [-ref] dat object}: <i>se</i> {3 sg [-ref] acc masc object}: <i>lo</i>					*	*

(56) a. *Se le dará el libro.*
 IMPERSONAL him/her give-3SG.FUTURE the book
 ‘Everybody will give him/her the book.’

b. *Juan_i se_i le_j presentó.*
 John REFLEXIVE him introduce-3SG.PAST
 ‘John introduced himself to him.’

Se also replaces *les*, the plural form of *le*, when *les* is adjacent to another clitic with an initial *l*. In Spanish, *los* marks the feature value set {3, pl, [-ref], acc, masc, object} and *las* marks the set {3, pl, [-ref], acc, fem, object} and *les* {3, pl, [-ref], object}. By positing the relevant realization constraints in (57), we can account for the

alternation of *les/los/las* and of *se/les*. See (58).

- (57) a. {3, pl, [-ref], acc, masc, object}: *los*: The feature value set {3, pl, [-ref], acc, masc, object} is realized by *los*. (LOS)
 b. {3, pl, [-ref], acc, fem, object}: *las*: The feature value set {3, pl, [-ref], acc, fem, object} is realized by *las*. (LAS)
 c. {3, pl, [-ref], object}: *les*: The feature value set {3, pl, [-ref], object} is realized by *les*. (LES)

(58) a. *Juan les dió el libro*. ‘John gave the book to **them**.’

{3 pl [-ref] dat object}	OCP <i>IV(C)</i>	*FS	LOS	LAS	LES	SE
a. {3 pl [-ref] dat object}: <i>los</i>					*!	*
b. {3 pl [-ref] dat object}: <i>las</i>					*!	*
☞ c. {3 pl [-ref] dat object}: <i>les</i>						*
d. {3 pl [-ref] dat object}: <i>se</i>					*!	
e. {3 pl [-ref] dat object}		*!***				

b. *Juan se los dió*. ‘John gave **them (masc)** to **them**.’

{3 pl [-ref] dat object} & {3 pl [-ref] acc masc object}	OCP <i>IV(C)</i>	*FS	LOS	LAS	LES	SE
a. {3 pl [-ref] dat object}: <i>les</i> {3 pl [-ref] acc masc object}: <i>los</i>	*!				*	*
b. {3 pl [-ref] dat object}: <i>se</i> {3 pl [-ref] acc masc object}: <i>les</i>			*!		*	*
☞ c. {3 pl [-ref] dat object}: <i>se</i> {3 pl [-ref] acc masc object}: <i>los</i>					*	*


The ranking OCP *IV(C)* >> LO, LOS, LA, LAS >> LE, LES >> SE shows another case of the emergence of the unmarked. The spurious *se* usually does not appear in the context of 3rd person because of higher-ranked realization constraints. However, *se* emerges to avoid a combination of two clitics with initial *l*'s.²⁸ It is hard to imagine how a morphologically restricted OT approach can account for the alternations of these clitics.

There are some relevant issues worthy of discussion. It is possible to formulate the OCP constraint at a morphosyntactic level, i.e., instead of saying that two clitics with initial *l*'s should not be adjacent, we can say that two 3rd person clitics cannot be adjacent so that *se* which does not realize 3rd person replaces one of the 3rd person clitics (Grimshaw 1997b, Neeleman and van de Koot 2005). Constraints like *{3}{3}, however, do not account for cases of OCP-triggered (allo)morph selection and haplology such as those in Hungarian (e.g., *olvas-ol* ‘you read’, **olvas-(o)sz*), Tswana (e.g., *N-ka-réka* ‘1sg-conditional mood-buy’, **ke-ka-réka*), and Mandarin

($le_{\text{PERFECTIVE-CRS}}$, $*le_{\text{PERFECTIVE}}$ le_{CRS}) in which the two adjacent morphs which are phonologically (partially) identical do not share a feature value.

Grimshaw 2001 proposes an alignment-based approach to the Spanish spurious *se* rule. Following Perlmutter 1971, she assumes that 3rd person clitics compete for a single templatic slot so that, for example, *le* and *lo* which realize 3rd person cannot co-occur. *Se* which realizes [+reflexive] is chosen by the grammar to replace *le*. She also assumes that both the input and output are comprised of phonologically unrealized morphosyntactic feature values. The grammar determines the winning feature bundle. Her approach is essentially a feature-changing mechanism similar to the feature impoverishment-plus-insertion mechanism in Distributed Morphology (e.g., Noyer 1998). Consider the following tableau. Candidate (a) (i.e., $*le\ lo$), according to Grimshaw, violates the constraint PERSON RIGHT because the feature value {3} appears twice in the output. Candidate (b) (i.e., *se lo*) which contains only one {3} wins.

(59) Tableau (adapted from Grimshaw 2001)

Input: {-ref 3 sg dat + -ref 3 sg acc}	PERSON RIGHT
a. {-ref 3 sg dat + -ref 3 sg acc}	*!
 b. {+ref -ref 3 sg acc}	

There are several problems with such a mechanism. Above all, as Grimshaw notices, the feature-changing mechanism in OT does not predict which feature (or feature combination) should be realized by which phonological form. We, therefore, have to refer to an extraordinary mechanism to spell out the morphosyntactic feature bundle. By contrast, without positing two separate and distinct mechanisms, we can get the phonological output which spells out the unrealized input feature value(s) through a single and unified mechanism on the basis of realization constraints. Second, it is not clear why we should assume *se* realizes [+ref] because *se* can also be an impersonal pronoun. Third, Grimshaw 2001 uses alignment constraints instead of OCP to account for the spurious *se* rule. However, alignment constraints such as PERSON RIGHT and CASE RIGHT do not account for cases of OCP-triggered haplology and allomorph selection in which adjacent phonologically (partially) identical morphs do not share a morphosyntactic feature value. Last but not least, it is not clear whatsoever why alignment constraints such as PERSON RIGHT which requires phonological forms realizing the feature of person to be produced or perceived later can actually work, given that there is no phonological form but abstract morphosyntactic feature value in the output.

2.4.5 Swedish haplology

In this section, I show that cases of haplology in Swedish are naturally accounted for under a realization OT approach while it is hard for a morphologically restricted OT approach to explain them without introducing special features. The Swedish data in question come from Stemberger 1981.

In Swedish, morphs with an identical shape but distinct morphemic content (or

meaning) behave differently with respect to haplology. Swedish has two *s*-suffixes that show haplology. The possessive *-s*²⁹ is affixed to all nouns (60a) except those ending in a sibilant (60b).

- (60) a. *Pers bil* ‘Per’s car’ (Stemberger 1981: 797)
 b. *Lars’ bil* ‘Lars’s car’

“The collective suffix *-s* occurs in only a few nouns, e.g. *höns* ‘chicken’, and in family names, where it shows haplology” (Stemberger 1981: 797):

- (61) a. *Perssons* ‘the Perssons’
 b. *Lorentz (tz = [s])* ‘(the) Lorentz(es)’

By contrast, the passive suffix *-(e)s* does not show haplology:

- (62) a. *bygg-* ‘build’, passive *bygges ~ byggs* (Stemberger 1981: 802)
 b. *lös-* ‘loosen’, passive *löses (*lös)*

Additionally, “[t]he present tense suffix *-[e]r* shows obligatory haplology after all stems that end in (always non-morphemic) /r/” (Stemberger 1981: 797):

- (63) a. *bygg-* ‘build’, pres. *bygger*
 b. *rör-* ‘move’, pres. *rör (*rörer)*

By contrast, the non-neuter pl(ural) suffix *-er* does not show haplology:

- (64) *kör* ‘choir’, pl *körer*

Moreover, the non-neuter def(inite) sg suffix *-en* may show haplology:

- (65) *orden* ‘order’, def sg *orden (*ordnen)* (Stemberger 1981: 798)

But the neuter definite plural suffix *-en* does not show haplology:

- (66) a. *våpen* ‘weapon’ def pl *våpnen (*vapen)*
 b. *hémman-an* ‘homestead’ def pl *hémmanen (*hemman)*
 c. *plúmm-on* ‘plum’ def pl *plúmmonen (*plummon)*

To account for the asymmetric behavior of affixes with an identical shape but different morphemic content, we can propose the relevant OCP constraint (67) and various realization constraints which may or may not outrank OCP (morph). In (68), all the affixes realized by the constraints outranking OCP (morph) need to be spelt out and therefore do not show haplology. By contrast, all the affixes realized by the constraints outranked by OCP (morph) show haplology.

(67) OCP (morph): A (*V*)C affix should not be adjacent to a stem with an identical final consonant.

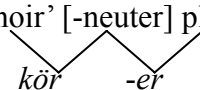
(68) Rankings

- a. {passive}: -(e)s >> OCP (morph) >> {poss}: -s, {plural}: -s
- b. {[-neuter], pl}: -er >> OCP (morph) >> {present}: -er³⁰
- c. {[+neuter], [+def], pl}: -en >> OCP (morph) >> {[-neuter], [+def], sg}: -en

We can use the ranking (68b) to demonstrate how our grammar works. Consider the tableaux in (69). In (69a), Candidate (b) is ruled out because the non-neuter plural marker *-er* is not spelt out. In (69b), Candidate (a) is ruled out because the presence of the present tense marker *-er* triggers a violation of OCP (morph).

(69) Tableaux

a. *kör* ‘choir’, pl

<i>kör</i> ‘choir’, [-neuter], pl	{[-neuter], pl}: -er	OCP (morph)	present: <i>-er</i>
☞ a. ‘choir’ [-neuter] pl 		*	
b. <i>kör</i> ‘choir’, [-neuter] pl	*!		

b. *rör-* ‘move’, present

<i>rör-</i> ‘move’, present	{[-neuter], pl}: -er	OCP (morph)	present: <i>-er</i>
a. ‘move’ present <i>rör- er</i>		*!	
☞ b. <i>rör</i> ‘move’, present			*

The asymmetric behavior of the affixes in question is readily formulated under a realization OT approach which assumes that both the input and output contain morphosyntactic feature values. By contrast, it is hard for a morphologically restricted OT approach to account for such asymmetric behavior without introducing morphosyntactic or diacritic features into the input and the grammar, which shows that phonological information such as “affix”, “root”, and “stem” does not constitute enough information to account for cases like Swedish haplology and we need to take meanings of affixes into consideration.

2.5 Problems for realization approaches and their solutions

In this section, I discuss problems caused by voicing assimilation for realization approaches and their possible solutions.

Yip 1998 accounts for the *s*-haplology in English (Stemberger 1981, Menn and MacWhinney 1984, among others) under a realization approach. To account for cases like *cat* + *-s*_{pl(ural)} + *-s*_{poss(essive)} → *cats*_{pl-poss} and *Katz* + *-s*_{pl} + *-s*_{poss} → *Katzes*_{pl-poss}, she

proposes the following constraints:

(70) Constraints (Yip 1998: 224)

- a. {pl}: -z: Plural is realized by the suffix -z.³¹
- b. {poss}: -z: Possessive is realized by the suffix -z.
- c. OCP (s): OCP (feature), where feature = [strident]
- d. DEP (V): Every vowel in the output has a correspondent in the input. (No vowel Insertion)

Yip assumes that input morphosyntactic feature values such as plural and possessive are phonologically unrealized and their corresponding phonological forms are determined by realization constraints.

Consider the tableaux in (71). The winning candidate in (71a) is *cat*-[z]. The question is how we account for the assimilation of /-z/ after *cat*. Notice that *cat*-[s] would violate the two realization constraints {pl}: -z and {poss}: -z.³²

(71) Tableaux (adapted from Yip 1998: 225)

a. *cat*, pl, poss

<i>cat</i> , pl, poss	{pl}: -z	{poss}: -z	OCP (s)	DEP (V)
<pre> pl poss a. <i>cat</i> -z -z </pre>			*!	
<pre> pl poss \ / b. <i>cat</i> -z </pre>				
<pre> pl poss c. <i>cat</i> -z -ə -z </pre>				*!

b. *Katz*, pl, poss

<i>Katz</i> , pl, poss	{pl}: -z	{poss}: -z	OCP (s)	DEP (V)
<pre> pl poss \ / a. <i>Katz</i> -z </pre>			*!	
<pre> pl poss \ / b. <i>Katz</i>, pl, poss </pre>	*!	*		
<pre> pl poss \ / c. <i>Katz</i> -ə -z </pre>				*

Even if we rank the constraint AGREE VOICE (Beckman 1998, Lombardi 1999, 2004) (72) higher than the constraints realizing -z, the grammar would pick up the incorrect output candidate *cat*-ə-z_{pl-poss}.

(72) AGREE VOICE: Obstruents in a cluster must agree in voicing.

Consider the tableau in (73). Candidate (a) is ruled out by AGREE VOICE. The correct output candidate *cat*-[s] loses in violation of both the realization constraints. The incorrect Candidate (c) wins despite its violation of DEP (V). Notice that DEP (V) cannot outrank the realization constraints because otherwise the grammar would choose the incorrect output candidate **Katz*_{pl-poss} or **Katz*-z_{pl-poss} (depending on the ranking of OCP and the constraints realizing -z) instead of the expected *Katz*-[ə]-z_{pl-poss}.

(73) Tableau

<i>cat</i> , pl, poss	AGREE VOICE	pl: -z	poss: -z	OCP (s)	DEP (V)
<p>a. <i>cat</i> -z</p>	*!				
<p>b. <i>cat</i> [s]</p>		*!	*		
<p>c. <i>cat</i> -ə -z</p>					*

There are several possible solutions. We could take a classical serial approach to this problem and assume that regular inflection applies before phonological processes such as voicing assimilation (Anderson 1992).³³ We could assume that certain types of outputs of regular inflection become inputs to phonology. In the case of English progressive assimilation, we could say that the output of regular inflection, for example, *cat*-[z] becomes an input to phonology and undergoes voicing assimilation. The cost of this approach would be to complicate our grammatical architecture and analytical processes. It might also cause questions such as which phenomenon belongs to which grammatical component in OT, a morphological domain or a phonological domain.

Another possibility would be to refer to underspecification of phonological features such as voice and assume that the plural marker or the possessive marker in English is an alveolar sibilant with an underspecified voicing feature [αvoice], i.e., the plural marker or the possessive marker could be either [s] or [z]. We could posit two types of realization constraints, one type realizing an alveolar sibilant with no specified voicing feature (74) and the other type realizing [z] which is the default.^{34,35}

- (74) a. {pl}: -s/-z: Plural should be realized as either -s or -z.
 b. {poss}: -s/-z: Possessive should be realized as either -s or -z.

See the tableau in (75). Notice that Candidate (e) is ruled out by AGREE VOICE. The winning candidate (f) satisfies the higher-ranked realization constraints.

(75) *cat*, pl, poss

<i>cat</i> , pl, poss	AGREE VOICE	pl: -s/-z	poss: -s/-z	OCP (s)	DEP (V)	pl: -z	poss: -z
a. <i>cat</i> , pl, poss		*!	*			*	*
b. <i>cat</i> <div style="display: inline-block; vertical-align: middle; margin-left: 20px;"> pl poss -z -z </div>	*!			*			
c. <i>cat</i> <div style="display: inline-block; vertical-align: middle; margin-left: 20px;"> pl poss -z -ə -z </div>	*!				*		
d. <i>cat</i> -ə <div style="display: inline-block; vertical-align: middle; margin-left: 20px;"> pl poss \ / -z </div>					*!		
e. <i>cat</i> <div style="display: inline-block; vertical-align: middle; margin-left: 20px;"> pl poss \ / -z </div>	*!						
f. <i>cat</i> <div style="display: inline-block; vertical-align: middle; margin-left: 20px;"> pl poss \ / -s </div>						*	*

This approach based on underspecification of phonological features is also not completely satisfactory. Above all, ranking constraints realizing phonologically less specific content higher than those realizing phonologically more specific content violates the specificity condition. It is also a redundant approach to posit two separate constraints realizing phonologically related content (i.e., -s/-z and -z).

The third solution to the problem caused by voicing assimilation would be to posit three separate constraints realizing the three English allomorphs [z], [s], and [əz] and specify the context where each allomorph occurs. See MacBride 2004 for such an approach. The cost would be to make these phonologically related allomorphs unrelated given that the phonological derivation of these allomorphs is sacrificed.

2.6 Conclusion

In this chapter, I examine an approach based on realization constraints (RC) in OT to (avoidance of) repetition of identical morphs. Our discussion focuses on both haplology and selection of phonologically unrelated allomorphs under the condition that the morphs in question are adjacent.³⁶ I compare the proposed realization OT approach to morphologically restricted OT approaches in accounting for both haplology and allomorph selection. I show that both approaches seem to be able to account for haplology. However, the realization approach readily accounts for OCP-triggered selection of phonologically unrelated (allo)morphs. I show that OCP-triggered (allo)morph selection in languages such as Greek, Hungarian, Spanish, and Tswana involve the emergence of a morphosyntactically less specific form. I propose the ranking $OCP \gg RC_{\text{specific}} \gg RC_{\text{less specific}}$ and show that such a ranking not only follows the specificity condition but also reflects the emergence of the

unmarked, i.e., a morphosyntactically less specific (allo)morph emerges to avoid a violation of OCP constraints though it generally does not show up in the context where a morphosyntactically more specific (allo)morph is expected to occur. By contrast, a morphologically restricted OT approach needs to stipulate in the input the relation between a morphosyntactically more specific (allo)morph and a morphosyntactically less specific one. It also needs to stipulate different input morphs in different morphosyntactic contexts. Such a stipulation will be avoided if we build certain morphemes into constraints. Morphosyntactic feature values are also necessary in accounting for Swedish haplology which varies with respect to meanings of affixes. Without extraordinary machinery, an approach based on the generation of morphosyntactic features is incapable of accounting for OCP-triggered haplology or (allo)morph selection whose nature is morphophonological, because such an approach lacks a systematic mechanism to spell out morphosyntactic features. Finally, I discuss problems caused by voicing assimilation for realization approaches and show that there are several potential solutions though all of them have defects, which shows that the relation between realization (morphology) and phonological processes awaits further research.

Notes

¹ Exceptions such as *sillily* and *surlily* are pointed out in a manuscript by Zwicky and Pullum, and quoted in Menn and MacWhinney 1984.

² Bonet 2004 compares the two OT approaches. A morphologically restricted OT approach is similar to the Morphs in the Input Hypothesis (MIH), i.e., all phonological forms are introduced through the input; a realization OT approach is similar to the Morphs through Constraints Hypothesis (MCH), i.e., some or all phonological forms are introduced through (realization) constraints. Her conclusion, however, is different from mine because she claims that a morphologically restricted OT approach is superior to a realization OT approach in that the latter does not successfully account for cases of voicing assimilation. We may observe comparative advantages of both the morphologically restricted OT model and the proposed realization OT model, i.e., the former may be more straightforward in accounting for phonological processes such as assimilation while the latter is better at explaining morphological phenomena such as blocking and extended exponece. See Chapter 4.

³ For various uses of the perfective aspectual marker *le*, see Li and Thompson 1981 for a very detailed discussion.

⁴ The context in which a CRS *le* occurs (i) indicates a changed state, (ii) corrects a wrong assumption, (iii) reports progress so far, (iv) determines what will happen next,

and (v) indicates the speaker's total contribution to the conversation at that point (Li and Thompson 1981: 244).

⁵ See also Neeleman and van de Koot 2005 for a similar case in Dutch. The function word *er* has several functions. It can be an expletive (X), the fronted argument of a preposition (P), a locative meaning 'there' (L), and a fronted noun associated with a numeral or a quantificational modifier (Q). They can co-occur when they are non-adjacent. Interestingly, when they are adjacent, only one *er* shows up.

(i) *Hebben er_{XLQ} [DP twee e] een auto gekocht?*
have there two a car bought
'Have two (of them) bought a car there?'

⁶ Notice that Yip 1998 does not clearly show that *le*, for example, realizes both [+perf] and CRS. Instead, she uses the representation "*mie_{PERF/CRS}-le*." However, some of her constraints such as Plural = *s* (the plural feature value corresponds to *s* in English) do seem to suggest that a morph should *realize* its corresponding morphosyntactic feature value(s). Additionally, representations such as those I formulate in Tableau (13) lead to an advantage, i.e., when we arrange morphemes in linear order, we can incorporate the semantic scope constraint (Bybee 1985, Rice 2000, Spencer 2003) into the grammar only if we know which morph realizes which semantic feature value.

⁷ This example is cited by Menn and MacWhinney from a manuscript by Zwicky and Pullum.

⁸ Golston 1995 argues that German function words such as *die* and *das* can be repeated because German syntax does not provide an alternative structure to avoid such repetition. He further argues that syntactic constraints leading to such repetition outrank (morpho)phonological ones against repetition, but it is not clear how such syntactic constraints are formulated.

⁹ See Mohanan and Mohanan 2003 which argues that OCP constraints are language-particular instantiations of the universal schema banning the adjacency of identical objects.

¹⁰ Brentari 1998 argues that the generalized OCP constraint is essentially what she calls the *REPEAT constraint. But the terminological difference between OCP and *REPEAT does not affect our analyses.

¹¹ De Lacy 1999 criticizes Plag's constraints with French data. E.g., *Baudis* + *-iste* [ist] → [bo.dist], *[bo.di.sist] (Corbin and Plénat 1992). *[bo.di.sist] does not clearly violate either OCP (ONSET) or OCP (ONSET, CODA) which disallows an onset to be identical to a coda within a syllable and thus poses a problem for Plag's account. If OCP (nucleus), which bans two adjacent identical nuclei, were violated, then the prediction would be: *Lenine* /Lenin/ + *iste* → [le.nist]. But the attested form is [leninist] not *[lenist].

¹² The French cases in (i) pose a potential problem for de Lacy's coalescence approach:

-
- (i) a. /ametist/ *amethyst* + /ist/ → [ametistist], *[ametist]
 b. /evaʁist/ *Évarist* {name} + /ist/ → [evaʁistist], *[evaʁist]

Under the coalescence approach, it is not clear why /ist/ coalesces with /analiz/ but not /ametist/. De Lacy says that the restriction banning complete coalescence of /ist/ is that “[a]n affix must have an output exponent that is unique to it.” It is not clear whether the constraint is universal given that many portmanteau morphs realize more than one morphosyntactic feature values cross-linguistically. Additionally, the data presented by de Lacy seem to point toward a case of morphologically conditioned truncation (Aronoff 1976) in which a specific shape of sound string (V(C) perhaps) is truncated (the underlined part is a truncated string) (ii). Just like in English, *demonstrable* is fine (*demonstrate* + *-able*) while **recognable* is ungrammatical (*recognize* + *-able* vs. *recogn-ition*), which shows that *ate* can be truncated while *ize* cannot, when *-able* is attached to the stem.

- (ii) a. /bodis/ *Baudis* {name} + /ist/ → [bodist], *[bodisist]
 b. /maʁini/ {name} + /ist/ → [maʁinist], *[maʁiniist]
 c. *Boulganine* + *-iste* → *boulganiste*, **boulganiniste*
 d. *Thucidide* + *-iste* → *thucidiste*, **thucididiste*
 e. *oasis* + *-ique* /ik/ → *oasique*, **oasisique*
 f. *stalactite* + *-ique* /ik/ → *stalactique*, **stalactitique*

¹³ See Section 2.4.2 for a further discussion and argument that a realization OT approach is superior to a morphologically restricted OT approach in accounting for the *-sz/-Vl* alternation in Hungarian.

¹⁴ The choice of *ou* and *ouk* may be determined by phonological constraints such as ONSET which requires every syllable to have an onset.

¹⁵ Golston 1995 also argues that an OCP-type constraint should outrank morphological constraints in the Greek case of selection of negative allomorphs. However, he does not propose any specific morphological constraints or discuss the details of realization.

¹⁶ Similarly, Halle 1997 proposes the Subset Principle in Distributed Morphology that a Vocabulary Item is subject to insertion if it contains all or a subset of grammatical features present in an input morpheme, and insertion does not take place if a Vocabulary Item contains grammatical features not present in an input morpheme.

¹⁷ The suffix *-sz* has several allomorphs. Subject to vowel harmony, *-(a)sz* and *-(e)sz* occur after stems ending in two consonants or a long vowel plus [t]; *-sz* occurs elsewhere (Rounds 2001: 26).

¹⁸ Carstairs (1988: 91, footnote 2) points out that “[i]n more conservative varieties, verbs of the minority ‘*ik*’ conjugation take only *-ol*, in all phonological contexts; in the innovative variety, however, all verbs conform to the pattern of the majority conjugation, with *-ol* and *(a)sz* distributed on a purely phonological basis.” The suffix *-Vl* does not realize the set {2nd person, sg, definite} which is consistently marked by *-Vd* (*-od* after a verbal stem and *-ad* after an inflectional suffix).

¹⁹ A constraint like (29a) or (29b) could be further parceled out in both a realization

constraint and an alignment constraint (McCarthy and Prince 1993, 2004, Russell 1997, Kager 1999, among others). For example, (29a) can be reformulated as two constraints in (i), but such a reformulation will not affect our analyses. Notice that it is not necessary to posit (ib) given the existence of (29d). In other words, (29a) can be replaced by (ia).

- (i) a. {2nd person, sg, indefinite, present, indicative}: *sz*: The feature value set {2nd person, sg, indefinite, present, indicative} is realized by the morph *sz*.
 b. ALIGN (Stem, R, {2nd person, sg, indefinite, present, indicative}, L): The right edge of a stem which is either lexical or inflected coincides with the left edge of the marker of {2nd person, sg, indefinite, present, indicative}.

²⁰ According to Cole, “Tswana has two distinct forms for the present tense principal positive, a long form [(SC-*a*-stem)] which employs the formative *-a-*, and a short form [(e.g., *ke-rêka* ... ‘I buy ..., I am buying ...’)] which has no such formative ... [T]he long form is used when the verb ends the sentence and has no objectival or adverbial adjunct following it, whereas the short form never occurs at the end of a sentence and therefore must always be followed by some adjunct” (p.243-244).

²¹ The prefix *ka-* in *ka-rêka* is a 1sg subject concord used in past subjunctive mood.

²² Specifically, the prefix *ba-* is the 3pl concord marker in agreement with a nominal subject or object which falls into Noun Class 1. “Class 1 contains only personal names” (Cole 1955: 70).

²³ I thank Eulàlia Bonet, Susana Huidobro, and Francesc Roca for providing and discussing the Spanish data.

²⁴ In some non-standard Spanish dialects, *lo* only marks non-human direct objects and therefore can only refer to *el libro* ‘the book’ but not *a Pedro* ‘to Peter’.

²⁵ In some non-standard Spanish dialects, *le* can also mark 3rd person sg non-reflexive *human* masculine direct objects:

- (i) *Juan le vió.*
 John him see-3rd sg past
 ‘John saw him.’ (*‘John saw it.’)

²⁶ See Maiden 2000 for various diachronic evidence for such a constraint.

²⁷ By contrast, Grimshaw 1997b claims that *se* marks nothing.

²⁸ In Standard Italian, *si* can be both a 3rd person reflexive and impersonal clitic. Interestingly, when a 3rd person reflexive clitic is expected to be adjacent to an impersonal clitic, we get *ci si*, **si si*. *Ci* can be both a 1st person plural and locative clitic. If *ci* is argued to be a default form, then the ranking OCP >> SI >> CI gives us another case of the emergence of the unmarked. If *si* is indeed a default form, we can use an output-to-output correspondence constraint which says a clitic copies the form of {1, pl}. This constraint is outranked by the equally specific constraint realizing *si*, which vacuously satisfies the specificity condition.

²⁹ The status of the Swedish possessive *-s* is controversial. See Norde 1997, 2001a, b which argue that the Swedish possessive *-s* is a clitic. See also Börjars 2003 for counterarguments. However, the status of the Swedish possessive *-s* is irrelevant to our discussion.

³⁰ The present singular suffix in Old Swedish was never haplogitized (Wessén 1965, cited in Stemberger 1981): *byggir, rörir*. This fact can be readily explained by the ranking {present sg}: *-ir* >> OCP (morph).

³¹ Yip uses the orthographic *s* instead of */-z/*, but I assume *s* corresponds to the traditional representation */-z/*.

³² Bonet 2004 discusses a similar problem caused by voicing assimilation in Catalan. She argues that positional faithfulness constraints are necessary in accounting for cases of regressive voicing assimilation in Catalan, so at least some phonological information needs to be in the input so that faithfulness constraints can operate. She further argues that a “radical” realization approach which assumes all phonological information is introduced through constraints rather than the input (Russell 1995) fails to account for cases of voicing neutralization in Catalan. The Catalan data discussed in Bonet 2004 are all lexemes, which seems to pose no problem for Yip’s 1998 approach in which the phonological information of lexemes is posited in the input.

³³ Russell 1997 also notices such a possibility.

³⁴ I thank Eric Baković for pointing out this solution to me.

³⁵ A similar analysis can be made of the alternation of the past tense allomorphs [t], [d], and [əd]. We can assume that a constraint realizing either [t] or [d] outranks a constraint realizing the default past tense marker [d].

³⁶ See Yip 1998 for a realization approach to avoidance of repetition of identical morphs which are non-adjacent (e.g., English disfavors two adjacent verbs both of which have the progressive aspect marker *-ing* (Ross 1972)) and avoidance of the context where such repetition is expected to occur.

Chapter Three

A realization OT approach to affix order in Lezgian

3.1 Introduction

In this chapter I continue to argue for the advantages of a realization OT model which associates morphosyntactic or semantic feature values with phonological forms through realization constraints. I show that this model readily incorporates universal generalizations about affix order. By contrast, rule-based morphological models such as Paradigm Function Morphology (Stump 2001) need to use extraordinary machinery to capture these universal generalizations. For most of the part, I use the Lezgian language to demonstrate how a realization OT model accounts for affix order.

This chapter is organized as follows. In section 3.2 I describe the relevant Lezgian data which come from Haspelmath (1993). I take a realization OT approach to the Lezgian data in section 3.3. I compare this approach with other alternative accounts in section 3.4 and conclude in section 3.5.

3.2 Lezgian

According to Haspelmath 1993, “Lezgian is spoken by about 400,000 people in southern Dagestan and northern Azerbaijan in the eastern Caucasus ... Lezgian morphology is overwhelmingly suffixing and agglutinating” (p.4). We focus our discussion on the data of the lowland Güne dialect, which is considered standard Lezgian. Lezgian has a rich inflectional system. We discuss both its nominal and verbal inflections.

3.2.1 Nominal inflection

According to Haspelmath 1993, Lezgian “[n]ouns are inflected for number (singular, plural), case, and localization (ad, sub, post, super, in)” (p.4). “There are eighteen cases in Lezgian: four grammatical ones (absolute, ergative, genitive, dative) and fourteen local cases divided into five localizations (ad, post, sub, super, in), each of which has three locatives (essive, elative, directive). One combination, the ‘in-directive’, is missing, so there are only $3 \times 5 - 1 = 14$ combinations” (p.74).

The following is an illustrative paradigm for singular forms. The root is *sew*. The suffix *-re* is an oblique suffix or ergative case marker. In (1) the suffix *-re* appears in every paradigmatic cell except the absolute. Notice that in the inelative form *sew-räj* ‘out of the bear’, the vowel *ä* [æ] coalesces the advancement feature [-back] of the vowel [e] of *-re* and the height feature [+low] of the vowel [a] of *-aj*. The genitive case marker is the suffix *-n*. The dative exponent is the suffix *-z*. The suffixes *-w*, *-q^h*, *-k*, and *-l* realize the localizations “ad”, “post”, “sub”, and “super”

respectively. The suffixes *-aj* and *-di* mark the locatives “elative” and “directive” respectively.

(1) Lezgian singular nominal paradigm: *sew* ‘bear’ (Haspelmath 1993: 74)

Absolutive	<i>sew</i>	‘the bear’
Ergative	<i>sew-re</i>	‘the bear’
Genitive	<i>sew-re-n</i>	‘of the bear’
Dative	<i>sew-re-z</i>	‘to the bear’
AdeSSive	<i>sew-re-w</i>	‘at the bear’
Adelative	<i>sew-re-w-aj</i>	‘from the bear’
Addirective	<i>sew-re-w-di</i>	‘toward the bear’
Postessive	<i>sew-re-q^h</i>	‘behind the bear’
Postelative	<i>sew-re-q^h-aj</i>	‘from behind the bear’
Postdirective	<i>sew-re-q^h-di</i>	‘to behind the bear’
Subessive	<i>sew-re-k</i>	‘under the bear’
Subelative	<i>sew-re-k-aj</i>	‘from under the bear’
Subdirective	<i>sew-re-k-di</i>	‘to under the bear’
Superessive	<i>sew-re-l</i>	‘on the bear’
Superelative	<i>sew-re-l-aj</i>	‘off the bear’
Superdirective	<i>sew-re-l-di</i>	‘onto the bear’
Inessive	<i>sew-re</i>	‘in the bear’
Inelative	<i>sew-räj</i>	‘out of the bear’

Let us go through the above-mentioned eighteen cases. The absolutive case form does not take an overt oblique suffix.

The ergative case form is identical to an absolutive case form plus an oblique suffix. There are ten oblique suffixes in Lezgian. See (2).

(2) Oblique suffixes in Lezgian (Haspelmath 1993: 74)

<i>-di</i>	<i>-a</i>	<i>-i</i>	<i>-u</i>	<i>-Adi</i>
<i>-rA</i>	<i>-Uni</i>	<i>-A</i>	<i>-U</i>	<i>-ci / -c’i / -či / -č’i / -ži</i>

Notice that some of the above suffixes contain capitalized *A* and *U* which stand for “harmonious vowels” that can harmonize with preceding root vowels. *A* stands for either /a/ (back, low) or /e/ (mid-front): /a/ appears in the environment of /a, u/; /e/ appears in the environment of /e, i, y, æ/. *U* stands for high vowels, either /u/, /i/, or /y/: /u/ appears in the environment of /a, u/; /i/ appears in the environment of /e, i, æ/; /y/ appears in the environment of /y/. See the table in (3) and the examples of *A*-suffix

and *U*-suffix in (4) and (5).

(3) Vowel Harmony alternations (Haspelmath 1993: 56)

A = /a/ in the environment of /a, u/
 = /e/ in the environment of /e, i, y, æ/

U = /u/ in the environment of /a, u/
 = /i/ in the environment of /e, i, (æ)/
 = /y/ in the environment of /y/

(4) *-rA* (an oblique suffix) has two allomorphs: *-ra* and *-re* (Haspelmath 1993: 57)

absolutive	ergative	
<i>lam</i>	<i>lam-ra</i>	‘donkey’
<i>č’ut</i>	<i>č’ut-ra</i>	‘flea’
<i>c’eh</i>	<i>c’eh-re</i>	‘goat’
<i>q’if</i>	<i>q’if-re</i>	‘mouse’
<i>ğüč’</i>	<i>ğüč’-re</i>	‘moth’

(5) *-Uni* (an oblique suffix) has three allomorphs: *-uni*, *-ini*, and *-üni* (Haspelmath 1993: 58)

absolutive	ergative	
<i>ğal</i>	<i>ğal-uni</i>	‘thread’
<i>č’ul</i>	<i>č’ul-uni</i>	‘belt’
<i>peš</i>	<i>peš-ini</i>	‘leaf’
<i>ric’</i>	<i>ric’-ini</i>	‘bowstring’
<i>q’ül</i>	<i>q’ül-üni</i>	‘foot; dance’

The distribution of six of the ten suffixes is predictable. The suffix *-di* is the default oblique suffix which attaches to “polysyllabic nouns (e.g. *bubá*, obl. *bubá-di* ‘father’)” and “monosyllabic words ending in a vowel, monosyllabic loanwords, and abbreviations” (Haspelmath 1993: 75).

The oblique suffix *-a* attaches to “[p]ersonal names ending in a consonant (e.g. *Farid*, obl. *Farid-a* ‘Farid’).” (Haspelmath 1993: 75)¹

The oblique suffix *-i* attaches to “[a]bstract nouns derived with *-wal* and Masdars (verbal nouns) in *-(u)n*” (Haspelmath 1993: 75). See (6).

(6) <i>jaru-wal</i>	obl.	<i>jarú-wil-i</i>	‘redness’
<i>k’él-un</i>	obl.	<i>k’él-un-i</i>	‘learning’
<i>q’í-n</i>	obl.	<i>q’í-n-i</i>	‘dying’
<i>fî-n-if</i>	obl.	<i>fî-n-if-i</i>	‘going’

The suffix *-i* also attaches to “all plural suffixes except *-bur*”² (Haspelmath 1993:

75). See (7).³

(7) <i>balk'an-ar</i>	obl. <i>balk'án-r-i</i>	'horses'
<i>buba-jar</i>	obl. <i>bubá-jr-i</i>	'fathers'
<i>dağ-lar</i>	obl. <i>dağ-lár-i</i>	'mountains'

The oblique suffix *-u* attaches to plurals in *-bur* (e.g. *jaru-bur*, obl. *jarú-bur-u* 'red ones').

The oblique suffix *-Adi* is used with "nouns that denote a non-discrete mass." See (8) (from Haspelmath 1993: 76).

(8) <i>nek</i>	obl. <i>nek'-édi</i>	'milk'
<i>c'ap</i>	obl. <i>c'ap-ádi</i>	'manure'

The oblique suffix "*-rA* is used with most native monosyllabic nouns that denote animals."⁴ See (9) (from Haspelmath 1993: 76).

(9) The oblique suffix *-rA* (= (4))

<i>lam</i>	<i>lam-ra</i>	'donkey'
<i>č'ut</i>	<i>č'ut-ra</i>	'flea'
<i>c'eh</i>	<i>c'eh-re</i>	'goat'
<i>q'if</i>	<i>q'if-re</i>	'mouse'
<i>ğüč'</i>	<i>ğüč'-re</i>	'moth'

"The distribution of the remaining four oblique stem suffixes [i.e. *-Uni*, *-A*, *-U*, *-ci/-c'i/-či/-č'i/-ži*]⁵ apparently does not follow from any semantic, morphological, or phonological principles. It has to be learned and remembered individually for each lexical item" (Haspelmath 1993: 77).

To briefly summarize, the formation of an ergative case form is complicated in that a nominal stem can combine with any one of ten oblique suffixes or ergative case markers, four of which are unproductive.

"The Genitive and Dative cases are formed by adding *-n* and *-z*, respectively, to the oblique stem [or ergative case form]. The localizations are formed from the oblique stem by adding their characteristic consonants *-w* (Ad), *-q^h* (Post), *-k* (Sub), respectively, and the Elative and Directive suffixes *-aj* and *-di* are added to the localization suffixes" (Haspelmath 1993: 74). See (10).

(10) Lezgian partial paradigm: *hül* ‘sea’ (adapted from Haspelmath 1993: 4)

Ergative	<i>hül-i</i>	‘the sea’
Genitive	<i>hül-i-n</i>	‘of the sea’
Dative	<i>hül-i-z</i>	‘to the sea’
Adelative	<i>hül-i-w-aj</i>	‘from the sea’
Addirective	<i>hül-i-w-di</i>	‘toward the sea’
Postelative	<i>hül-i-q^h-aj</i>	‘from behind the sea’
Postdirective	<i>hül-i-q^h-di</i>	‘to behind the sea’
Subelative	<i>hül-i-k-aj</i>	‘from under the sea’
Subdirective	<i>hül-i-k-di</i>	‘to under the sea’
Superessive	<i>hül-é-l</i>	‘on the sea’
Superrelative	<i>hül-é-l-aj</i>	‘from on the sea’
Superdirective	<i>hül-é-l-di</i>	‘onto the sea’
Inessive	<i>hül-é</i>	‘in the sea’
Inelative	<i>hül-áj</i>	‘from in the sea’

The formation of the “in” and “super” localization is more complicated. The in localization does not have an exponent consonant. It is formed by lowering the final vowel of the oblique stem: “[S]tressed *-ú, -ú̃, -í* become *-á, -é, -é*, while unstressed *-u* and *-i* both become *-a*⁶... If the final vowel of the oblique stem is [not high], the Inessive is identical to the oblique stem (and thereby to the Ergative case)” (Haspelmath 1993: 78). The inelative is built on the inessive form by adding *-aj*: If the inessive ends in *-a*, then the inelative ends in *-aj*; if the inessive ends in *-e*, then the inelative ends in *-ä[æ]j*. See (11).

(11) The formation of the inessive and inelative (the oblique suffix separated by a period, the lowered vowel of the inessive separated by a hyphen) (Haspelmath 1993: 78)

Oblique stem	Inessive	Inelative	
<i>čarx.ú</i>	<i>čarx-á</i>	<i>čarx-áj</i>	‘rock’
<i>q̃ül.ú⁷</i>	<i>q̃ül-é</i>	<i>q̃ül-áj</i>	‘wheat’
<i>čar.čí</i>	<i>čar.č-é</i>	<i>čar.č-áj</i>	‘paper’
<i>q̃acú-bur.u</i>	<i>q̃acú-bur-a</i>	<i>q̃acú-bur-aj</i>	‘green ones’
<i>Afrika.di</i>	<i>Afrika.d-a</i>	<i>Afrika.d-aj</i>	‘Africa’
<i>Murád.a</i>	<i>Murád-a</i>	<i>Murád-aj</i>	‘Murad’
<i>č’iž.r.é</i>	<i>č’iž.r-é</i>	<i>č’iž.r-áj</i>	‘bee’

The exponent consonant of the super localization is *-l*. The superessive is built on the form of the inessive by adding the suffix *-l*. “The Superdirective is formed completely regularly from the Superessive by adding the Directive suffix *-di*” (Haspelmath 1993: 79). The formation of the superrelative is more complicated. The superrelative is derived by combining *-l* and *-aj* (super: *-l*, elative: *-aj*) with either an inessive form whose oblique suffix ends in a stressed vowel or an ergative form whose oblique suffix ends in an unstressed vowel. See (12).

(12) Oblique stem	Inessive	Superrelative	
<i>čarx.ú</i>	<i>čarx-á</i>	<i>čarx.á-l-aj</i>	‘rock’
<i>q̃ül.ú</i>	<i>q̃ül-é</i>	<i>q̃ül.é-l-aj</i>	‘wheat’
<i>čar.čí</i>	<i>čar.č-é</i>	<i>čar.čé-l-aj</i>	‘paper’
<i>q̃acú-bur.u</i>	<i>q̃acú-bur-a</i>	<i>q̃acú-bur.u-l-aj</i>	‘green ones’
<i>Afrika.di</i>	<i>Afrika.d-a</i>	<i>Afrika.di-l-aj</i>	‘Africa’

To briefly summarize, the absolutive case form does not take an oblique suffix. The ergative case form contains an oblique suffix, which occurs in every paradigmatic cell except the absolutive. The inessive case form is derived by lowering the final vowel of an oblique suffix if the vowel is high. The inessive form underlies the in localizations, superessive and superdirective forms. Either the inessive form or the ergative form underlies the superrelative depending on whether the final vowel of the oblique suffix is stressed or not.

Next, let us consider restrictions on the order of Lezgian exponent affixes. The generalization is that localization markers are closer to the nominal stem than locative (essive, elative, directive) markers. The suffixes *-w* (ad), *-q^h* (post), *-k* (sub) and *-l* (super) precede the suffixes *-aj* (elative) and *-di* (directive). It is not obvious how the zero suffix of the essive is ordered with the localization markers, but I assume that it follows the generalization. The in localization does not have its characteristic consonant and there is no in-directive in Lezgian. I assume that the inelative form which is derived by combining the elative marker *-aj* with the inessive form does not contradict our generalization. See (13).

(13) Lezgian partial nominal paradigm (*sew* ‘bear’)

AdeSSive	<i>sew-re-w</i>	‘at the bear’
Adelative	<i>sew-re-w-aj</i>	‘from the bear’
Addirective	<i>sew-re-w-di</i>	‘toward the bear’
Postessive	<i>sew-re-q^h</i>	‘behind the bear’
Postelative	<i>sew-re-q^h-aj</i>	‘from behind the bear’
Postdirective	<i>sew-re-q^h-di</i>	‘to behind the bear’
Subessive	<i>sew-re-k</i>	‘under the bear’
Subelative	<i>sew-re-k-aj</i>	‘from under the bear’
Subdirective	<i>sew-re-k-di</i>	‘to under the bear’
Superessive	<i>sew-re-l</i>	‘on the bear’
Superelative	<i>sew-re-l-aj</i>	‘off the bear’
Superdirective	<i>sew-re-l-di</i>	‘onto the bear’
Inessive	<i>sew-re</i>	‘in the bear’
Inelative	<i>sew-räj</i>	‘out of the bear’

A further generalization can be made from the paradigm in (13). We can assume that the locatives and localizations roughly denote the following meanings:

(14)

Essive:	‘in a position of’	Ø
Elative:	‘from a position of’	-aj
Directive:	‘toward a position of’	-di
Ad:	‘nearby’	-w
Post:	‘behind’	-q ^h
Sub:	‘under’	-k
Super:	‘on’	-l
In:	‘in’	(lowering of a final high vowel)

The paradigm in (13) leads to the generalization that the locatives scoping over the localizations corresponds to the order in which the locative markers are farther away from the nominal stem than the localization markers (e.g. subelative: *sew-re-k-aj* ‘**from** under the bear’).

This generalization seems to hold in an earlier stage of the Lezgian language. According to Haspelmath 1993, except for the subdirective case which “[d]espite its name ... never expresses the locative notion ‘direction toward below’” (p.98), all the other 13 cases originally express the meanings in (13). They thus show semantic compositionality and denote “scope”.

This generalization is, however, complicated by the fact that most cases in (13) are either rarely used in the modern standard language or used in functions different from those in (13). For example, the adelative case “originally expresses movement

away from the location ‘near, by’, but it is now mostly used in a more abstract sense” (Haspelmath 1993: 90). See (15) (from Haspelmath 1993: 91).

- (15) *I Müškür xalu.di-z ča-waj wuč k’an-zawa-t’a?*
 this Müškür uncle-DAT we-ADEL what:ABS want-IMPF-CND
 ‘I wonder what this Müškür-xalu wants from us?’

The exceptions are the superessive ‘on’, superrelative ‘from on’, inessive ‘in’, and inelative ‘from in’. See (16).

- (16) a. *Č’ur.a-l wad jac amuq’-na.*
 pasture-SUPERESSIVE five ox stay-AORIST
 ‘Five oxen were still on the pasture.’ (Haspelmath 1993: 98)
- b. *Nurali buba balk’an.di-laj ewič’-na.*
 Nurali father horse-SUPERRELATIVE descend-AORIST
 ‘Father Nurali got off the horse.’ (Haspelmath 1993: 100)
- c. *Pahliwan-ar isätda či xür-e awa.*
 artist-PLURAL now we: GEN village-INESSIVE be.in
 ‘The tightrope walkers are now in our village.’ (Haspelmath 1993: 102)
- d. *Xatimat.a gičin.d-aj nek ca-zwa-j.*
 Xatimat(ERG) jug-INELATIVE milk pour-IMPERFECTIVE-PAST
 ‘Xatimat was pouring milk from a jug.’ (Haspelmath 1993: 103)

The generalizations about case forms also apply to plurals. “Generally, the plural suffix is the stress-attracting suffix *-Ar* or its stress-neutral variant *-ar*. The default plural suffix is stress-neutral *-ar*. Almost all polysyllabic nouns form their plural in *-ar*” (Haspelmath 1993: 71). See (17).

- (17) *muhmán-ar* ‘guests’, *penžér-ar* ‘windows’

Nouns ending in a vowel form their plural in *-jar* (e.g. *didé-jar* ‘mothers’). “Most monosyllabic nouns that end in a consonant form their plural in *-Ar*” (Haspelmath: 71).⁸ See (18).

- (18) *tar-ár* ‘trees’, *ğül-ér* ‘husbands’

Consider the paradigm in (19). The root is *hül*. Let us first look at the singulars. The absolutive form *hül* does not have the oblique suffix *-í*. The ergative form *hülí* has an oblique suffix *-í*, which appears everywhere except in the absolutive. The inessive form *hülé* is derived by lowering the vowel [í] of the oblique suffix. The inessive form

underlies the in and super localizations. Notice that the ergative form *hūli* should be listed because the distribution of the oblique suffix *-i* which combines with a few non-derived monosyllabic nouns is unpredictable.

Next, let us look at the plurals. The plural marker is *-ér*. The absolutive plural form *hūlér* does not contain the oblique suffix *-i*. The ergative plural form *hūléri* contains the oblique suffix *-i*. The inessive plural form *hūléra* is derived by lowering the vowel [i] of the oblique suffix. The inessive plural form underlies the in localizations and the superessive and superdirective forms. Notice that the distribution of the unstressed *-i* is predictable in that it attaches to “all plural suffixes except *-bur*” (Haspelmath 1993: 75). The oblique suffix vowel [i] of the superrelative plural form *hūl-ér-i-l-aj* is identical to that of the ergative form (*hūl-ér-i*) because it is unstressed.

(19) Lezgian nominal paradigm (*hūl* ‘sea’) (adapted from Haspelmath 1993: 4)

	Singular	Plural
Absolutive	<i>hūl</i>	<i>hūl-ér</i>
Ergative	<i>hūl-i</i>	<i>hūl-ér-i</i>
Genitive	<i>hūl-i-n</i>	<i>hūl-ér-i-n</i>
Dative	<i>hūl-i-z</i>	<i>hūl-ér-i-z</i>
Adessive	<i>hūl-i-w</i>	<i>hūl-ér-i-w</i>
Adelative	<i>hūl-i-w-aj</i>	<i>hūl-ér-i-w-aj</i>
Addirective	<i>hūl-i-w-di</i>	<i>hūl-ér-i-w-di</i>
Subessive	<i>hūl-i-k</i>	<i>hūl-ér-i-k</i>
Subelative	<i>hūl-i-k-aj</i>	<i>hūl-ér-i-k-aj</i>
Subdirective	<i>hūl-i-k-di</i>	<i>hūl-ér-i-k-di</i>
Postessive	<i>hūl-i-q^h</i>	<i>hūl-ér-i-q^h</i>
Postelative	<i>hūl-i-q^h-aj</i>	<i>hūl-ér-i-q^h-aj</i>
Postdirective	<i>hūl-i-q^h-di</i>	<i>hūl-ér-i-q^h-di</i>
Superessive	<i>hūl-é-l</i>	<i>hūl-ér-a-l</i>
Superrelative	<i>hūl-é-l-aj</i>	<i>hūl-ér-i-l-aj</i>
Superdirective	<i>hūl-é-l-di</i>	<i>hūl-ér-a-l-di</i>
Inessive	<i>hūl-é</i>	<i>hūl-ér-a</i>
Inelative	<i>hūl-áj</i>	<i>hūl-ér-aj</i>

As we can see from the paradigm in (19), all case markers are farther away from the nominal stem *hūl-* than the plural marker *-ér*. The order of the locative and localization markers is consistent in both singulars and plurals, i.e., the locative markers are farther away from the nominal stem than the localization markers.

To summarize the patterns about nominal inflection in Lezgian, the absolutive

form does not have an oblique suffix while the ergative form has an oblique suffix, which occurs everywhere except in the absolutive. The inessive form which underlies the in localizations and the superessive and superdirective forms is derived by lowering the final vowel of the oblique suffix if the vowel is high. The ergative or inessive form underlies the superrelative form depending on whether the final vowel of the oblique suffix is stressed or not. All case markers are farther away from the nominal stem than number markers. The locatives scope over the localizations, which corresponds to the order in which the locative markers are farther away from the nominal stem than the localization markers. The generalization about the order of the locative and localization markers with respect to semantic scope generally holds in an earlier stage of the Lezgian language and remains visible in some of the case forms in the modern standard language.

3.2.2 Verbal inflection

According to Haspelmath 1993, Lezgian has “two morphological verb classes,” strong verbs and weak verbs. Strong verbs always take stressed thematic vowels when inflected while weak verbs do not combine with thematic vowels. The citation form of a verb is the masdar (verbal noun). Each strong verb has three types of stems, the masdar stem underlying the masdar, optative, etc., the imperfective stem underlying the infinitive, imperfective, future, prohibitive, etc., the aorist stem underlying the aorist, perfect, aorist participle, etc. A weak verb only has one verbal stem.

Consider the paradigm in (20). The verbs *raxun* ‘talk’ and *fin* ‘go’ are strong verbs and the verb *kisun* ‘fall asleep’ is a weak verb. The verb *raxun* ‘talk’ has three types of verbal stems. The masdar stem *rax-ú-* contains the thematic vowel [ú]. Both the imperfective and aorist stem *rax-á-* contains the thematic vowel [á]. The verb *fin* ‘go’ also has three types of verbal stems. Both the masdar and imperfective stem *f-i-* contains the thematic vowel [í]. The aorist stem *f-é-* contains the thematic vowel [é]. The weak verb *kisun* ‘fall asleep’ has only one verbal stem *kis-*.

(20) Lezgian partial verbal paradigm (Haspelmath 1993: 122)⁹

	strong verb	strong verb	weak verb
	<i>raxun</i> ‘talk’	<i>fin</i> ‘go’	<i>kisun</i> ‘fall asleep’
base	<i>rax-</i>	<i>f-</i>	<i>kis-</i>
Masdar stem	<i>rax-ú-</i>	<i>f-i</i>	<i>kis-</i>
Masdar	<i>rax-ú-n</i>	<i>f-i-n</i>	<i>kis-un</i>
Optative	<i>rax-ú-raj</i>	<i>f-í-raj</i>	<i>kis-raj</i>
Imperfective stem	<i>rax-á-</i>	<i>f-i-</i>	<i>kis-</i>
Infinitive	<i>rax-á-z</i>	<i>f-i-z</i>	<i>kis-iz</i>
Imperfective	<i>rax-á-zwa</i>	<i>f-í-zwa</i>	<i>kis-zawa</i>
Future	<i>rax-á-da</i>	<i>f-í-da</i>	<i>kis-da</i>
Prohibitive	<i>rax-á-mir</i>	<i>f-í-mir</i>	<i>kis-mir</i>
Aorist stem	<i>rax-á-</i>	<i>f-e-</i>	<i>kis-</i>
Aorist	<i>rax-á-na</i>	<i>f-é-na</i>	<i>kis-na</i>
Perfect	<i>rax-á-nwa</i>	<i>f-é-nwa</i>	<i>kis-nawa</i>
Aorist participle	<i>rax-á-j</i>	<i>f-e-jí</i>	<i>kis-aj</i>

The distribution of thematic vowels is unpredictable¹⁰ and perhaps determined by each verbal root, so the three types of verbal stems of a strong verb have to be listed. Consider (21). It is unpredictable that, for example, the verbal root *ac’-* ‘be filled’ combines with the thematic vowel [á] in its aorist stem while the verbal root *qač-* ‘take’ combines with the thematic vowel [ú] in its aorist stem, although both roots take the thematic vowel [ú] in their masdar and infinitive stems. Haspelmath (1993: 123-125) lists dozens of strong verbs with their corresponding combinations of thematic vowels. It seems that each verbal root has its own combination of thematic vowels which falls into a pattern of vowel combination in (21).

(21) Some Lezgian strong verbs (Haspelmath 1993: 123)

Thematic vowels	Masdar	Infinitive	Aorist	
<i>U-U-A</i>	<i>ac’-ú-n</i>	<i>ac’-ú-z</i>	<i>ac’-á-na</i>	‘be filled’
	<i>q^hič-í-n</i>	<i>q^hič-í-z</i>	<i>q^hič-é-na</i>	‘sprinkle’
<i>U-A-A</i>	<i>gat-ú-n</i>	<i>gat-á-z</i>	<i>gat-á-na</i>	‘beat’
	<i>q^hür-ú-n</i>	<i>q^hür-é-z</i>	<i>q^hür-é-na</i>	‘laugh’
<i>U-A-U</i>	<i>č’ug-ú-n</i>	<i>č’ugw-á-z</i>	<i>č’ug-ú-na</i>	‘pull, draw’
<i>U-U-U</i>	<i>qač’-ú-n</i>	<i>qač’-ú-z</i>	<i>qač’-ú-na</i>	‘take’
<i>u-i-a</i>	<i>ak’-u-n</i>	<i>ek’-i-z</i>	<i>ak’-a-na</i>	‘stick, get stuck’

Let us consider a more complicated paradigm (22). The inflected forms of the verb *fin* ‘go’ are placed in three categories based on the three verbal stems of a strong verb. The masdar exponent of strong verbs is the suffix *-n*. The optative exponent of both strong and weak verbs is the suffix *-raj*. The imperative form of the verb *fin* is

the suppletive form *alad*. The infinitive exponent of strong verbs is the suffix *-z*. The imperfective marker of strong verbs is the suffix *-zwa*. The continuative imperfective exponent of strong verbs is the suffix *-zma*. The future exponent of both strong and weak verbs is the suffix *-da*. The hortative marker of strong verbs is the suffix *-n*. The prohibitive exponent of both strong and weak verbs is the suffix *-mir*. The posterior converb marker of both strong and weak verbs is the suffix *-daldi*. The graduative converb marker of strong verbs is the suffix *-rdawaj*. The aorist marker of both strong and weak verbs is the suffix *-na*. The perfect marker of strong verbs is the suffix *-nwa*. The continuative perfect marker of strong verbs is the suffix *-nma*. The aorist converb exponent is “invariably *-na*” (Haspelmath 1993: 131). The past tense marker is either *-j* or *-ir*. The past tense suffix *-ir* follows the negative marker *-č*. The participle marker is the suffix *-j*. Notice that the aorist participle marker of the verb *fin* is the suffix *-ji*.¹¹ The aorist participle marker in the negative environment (i.e. *te-fe-j*) is the suffix *-j*. The negative marker is either the suffix *-č* or the prefix *tA-* (*te-* in this case). The imperative cannot be negated and we can simply assume that a negated imperative is a prohibitive.

(22) Primary verb forms (*fin* ‘go’) (adapted from Haspelmath 1993: 127)

	affirmative	negative	affirmative participle	negative participle
Masdar	<i>fi-n</i>	<i>te-fi-n</i>		
Optative	<i>fi-raj</i>	<i>te-fi-raj</i>		
Imperative	<i>alad</i>	---		
Infinitive	<i>fi-z</i>	<i>te-fi-z</i>		
Imperfective	<i>fi-zwa</i>	<i>fi-zwa-č</i>	<i>fi-zwa-j</i>	<i>te-fi-zwa-j</i>
Past Imperfective	<i>fi-zwa-j</i>	<i>fi-zwa-č-ir</i>		
Continuative Imperfective	<i>fi-zma</i>	<i>fi-zma-č</i>	<i>fi-zma-j</i>	<i>te-fi-zma-j</i>
Past Cont. Imperfective	<i>fi-zma-j</i>	<i>fi-zma-č-ir</i>		
Future	<i>fi-da</i>	<i>fi-da-č</i>	<i>fi-da-j</i>	<i>te-fi-da-j</i>
Past Future	<i>fi-da-j</i>	<i>fi-da-č-ir</i>		
Hortative	<i>fi-n</i>	<i>te-fi-n</i>		
Prohibitive	<i>fi-mir</i>	---		
Posterior converb	<i>fi-daldi</i>	---		
Graduative converb	<i>fi-rdawaj</i>	---		
Aorist	<i>fe-na</i>	<i>fe-na-č</i>	<i>fe-ji</i>	<i>te-fe-j</i>
Past Aorist	<i>fe-na-j</i>	<i>fe-na-č-ir</i>		
Perfect	<i>fe-nwa</i>	<i>fe-nwa-č</i>	<i>fe-nwa-j</i>	<i>te-fe-nwa-j</i>
Past Perfect	<i>fe-nwa-j</i>	<i>fe-nwa-č-ir</i>		
Continuative Perfect	<i>fe-nma</i>	<i>fe-nma-č</i>	<i>fe-nma-j</i>	<i>te-fe-nma-j</i>
Past Continuative Perfect	<i>fe-nma-j</i>	<i>fe-nma-č-ir</i>		
Aorist converb	<i>fe-na</i>	<i>te-fe-na</i>		

The exponents in (22) are listed in (23).

(23) Inflectional exponents of *fin* ‘go’

masdar: <i>-n</i>	optative: <i>-raj</i>	imperative: <i>alad</i>
infinitive: <i>-z</i>	imperfective: <i>-zwa</i>	cont. imperfective: <i>-zma</i>
future: <i>-da</i>	past: <i>-j/-ir</i>	hortative: <i>-n</i>
prohibitive: <i>-mir</i>	posterior converb: <i>-daldi</i>	grad. converb: <i>rdawaj</i>
aorist: <i>-na</i>	perfect: <i>-nwa</i>	cont. perfect: <i>-nma</i>
aorist converb: <i>-na</i>	negative: <i>-č/tA-</i>	participle: <i>-j</i>
aorist participle: <i>-ji</i>		

There are basically three types of negative markers in Lezgian, the suffix *-č*, the prefix *t(A)-*, and periphrastic prefixal negative forms. We can say that “the participles, the converbs, the Infinitive, the Masdar, and the Periphrasis forms are non-finite, and that the remaining verb forms are finite. Within the group of finite verb forms, the Hortative, the Optative, the Imperative, and the Prohibitive will be said to be non-indicative, the others are indicative” Haspelmath (1993: 127). The suffix *-č* negates indicatives such as imperfectives and past imperfectives. If the suffix *-č* co-occurs with past tense forms, it needs to precede the past tense marker *-ir*.

The prefix *t(A)-* is used in some strong verbs to negate non-indicatives (e.g. *awun* (masdar) ‘do’, *t-awun* (negated masdar); *q’un* (masdar) ‘hold’, *ta-q’un* (negated masdar)).¹² Moor 1985 lists 18 such strong verbs which are cited in Haspelmath 1993. “In Uslar (1896: §258, §274), the class is much larger. Uslar lists about sixty verbs with inflectional prefixal negation” (Haspelmath 1993: 133).

Most verbs take periphrastic prefixal negative forms which only apply to non-indicatives and “are formed with the auxiliary *t-awun* ‘not do’ and the Periphrasis form. The Periphrasis form is always identical to the base in weak verbs, and in strong verbs it is most commonly identical to the Masdar” (Haspelmath 1993: 133). See (24).

(24) Illustrative partial paradigms (adapted from Haspelmath 1993: 135)

	<i>katun</i> ‘run’ (weak verb)		<i>raxun</i> ‘talk’ (strong verb)	
	affirmative	negative	affirmative	negative
Masdar	<i>kát-un</i>	<i>kat tawú-n</i>	<i>raxú-n</i>	<i>raxú-n tawú-n</i>
Infinitive	<i>kát-iz</i>	<i>kat tiji-z</i>	<i>raxá-z</i>	<i>raxá-n tiji-z</i>
Impf. participle	<i>kát-zawa-j</i>	<i>kat tiji-zwa-j</i>	<i>raxá-zwa-j</i>	<i>raxá-n tiji-zwa-j</i>
	<i>awun</i> ‘do’			
	affirmative	negative		
Masdar	<i>awú-n</i>	<i>t-awú-n</i>		
Infinitive	<i>iji-z</i>	<i>t-iji-z</i>		
Impf. participle	<i>iji-zwa-j</i>	<i>t-iji-zwa-j</i>		

As we can see from (22), the past tense markers *-j*, *-ir* are farther away from the verbal stem than the tense-aspect markers *-zwa* (imperfective), *-zma* (continuative imperfective), *-da* (future), *-na* (aorist), *-nwa* (perfect), and *-nma* (continuative perfect). This order also applies to weak verbs which take the imperfective marker *-zawa*, the continuative imperfective marker *-zama*, the perfect marker *-nawa*, and the continuative perfect marker *-nama*.

Let us go through the above tense-aspect forms. “The Imperfective typically refers to progressive situations, i.e. processes going on at the time of reference” (Haspelmath 1993: 140). See (25).

- (25) *Farida, Farida, wuna ana wuč iji-zwa?*
 Farida Farida you:ERG there what:ABS do-IMPERFECTIVE
 ‘Farida, Farida, what are you doing there?’ (Haspelmath 1993: 140)

Despite its name, the future not only refers to future situations, but also to habitual situations (26).¹³

- (26) a. *Širwan.di-z fi-da-j-bur sad=q’we juq̃.u-z*
 [Širwan-DATIVE go-FUTURE-PTP]-SBST.PL one=two day-DATIVE

Baku.d-a amuq’-da-j.
 Baku-INESSIVE stay-FUTURE-PAST
 ‘Those who went to Širwan used to stay one or two days in Baku.’
 (Haspelmath 1993: 141)

- b. *Gzaf q’in q’a-da-j-da gzaf tab-ni iji-da.*
 [much oath take-FUTURE-PTP]-SBST.SG(ERG) much lie-also do-FUTURE
 ‘He who swears a lot also lies a lot.’ (Haspelmath 1993: 141)

The future can be considered an aspect when it expresses habitual situations (cf. Bybee 1985). “The Past is only compatible with the habitual meaning of the Future, not with the future meaning” (Haspelmath 1993: 142).

The Aorist refers to “perfective events in the past” (Haspelmath 1993: 142):

- (27) *Sadwil.i wa aq’ulluwil.i abur q̃utarmiš-na.*
 unity(ERG) and cleverness(ERG) they save-AORIST
 ‘Unity and cleverness saved them.’ (Haspelmath 1993: 142)

“The Past Aorist refers to situations in the remote past, situations that took place before the main story line, situations that do not obtain anymore, and situations whose effect has been canceled” (Haspelmath 1993: 143).

The perfect refers to “past events with current relevance [(28a)], including events that are presented as ‘hot news’ [(28b)]” (Haspelmath 1993: 143).

(28) a. *Zi pul na aq̃ud-nawa!*
 I:GEN money you:ERG take.away-PERFECT
 ‘You have stolen my money!’ [the theft has just occurred]
 (Haspelmath 1993: 143)

b. *Kwe-z telegramma ata-nwa!*
 you.all-DAT telegram come-PERFECT
 ‘A telegram has come for you-all!’ (Haspelmath 1993: 144)

The past perfect “expresses temporal precedence (anteriority) to another past situation” (Haspelmath 1993: 145).

“The Continuative, combined either with the Imperfective or with the Perfect, adds the semantic element ‘still’” (Haspelmath 1993: 145).

Participles express relative clauses in Lezgian. “The various tense-aspect forms of the participles generally have the same temporal-aspectual meaning as the corresponding finite forms. Future participles [(FUT-PTP)] may have future [(29a)] or habitual [(29b)] meaning” (Haspelmath 1993: 155).

(29)a. *A xwanaxwa.di-z q̃e za koncert.d-a ja-da-j*
 that friend-DATIVE [[today I:ERG concert-INCESS play-FUT(URE)-PTP]

daldam xutax-iz k’an-zawa.
 drum take.away-INFINITIVE] want-IMPERFECTIVE
 ‘That friend wants to take away the drum that I will play today at the concert.’ (Haspelmath 1993: 155)

b. *Am ja marf, ja gar, ja cif te-fi-da-j,*
 it: ABS [or rain or wind or fog NEG-go-FUT-PTP]

q̃wan.ci-n alamat.di-n q̃ele tir.
 stone-GEN wonder-GEN fortress COPULA:PAST
 ‘It was a wonderful stone fortress in which neither rain, nor wind, nor fog entered.’ (Haspelmath 1993: 156)

“Continuative participles have continuative meaning” (Haspelmath 1993: 156):

(30) *Dide.di sufra ek’ä-na, axpa ada-l*
 mother(ERG) cloth spread-AORIST then it-SUPERESSIVE

hele rga-zma-j samovar ecig-na.
 [still boil-IMPF.CONT-PTP] samovar put-AORIST
 ‘Mother spread out a cloth, and then she put a samovar on it that was still boiling.’

“Perfect participles have perfect meaning” (Haspelmath 1993: 156):

- (31) *Q^hen-ar* *c'ra-na* *na luhudi, abur* *ife-nwa-j*
 shadow-PL dissolve-AORIST as- -if they [[heat-PERFECT-PTP]
- pič.ina-l* *ecig-nawa-j* *murk'uc'-ar* *tir.*
 oven-SUPERESSIVE put-PERFECT-PTP] icicle-PL COPULA:PAST
- ‘The shadows dissolved as if they were icicles which had been put on a hot oven.’

“The Aorist participle has past meaning” (Haspelmath 1993: 156):

- (32) *Q'ara.di-z* *awat-aj* *q̃izil* *q'alu* *že-da-č.*
 [mud-DAT fall-AORIST PTP] gold dirty become-FUT-NEG
- ‘Gold which has fallen into the mud does not become dirty.’

The participle is a relative clause marker just like English *which*, *that*, *who*, etc. and scopes over temporal-aspectual markers within a relative clause, which corresponds to the order in which the participle is farther away from the verb stem than temporal-aspectual markers.

To summarize the patterns about verbal inflection, Lezgian has both strong and weak verbs. A strong verb usually has three types of stems: the masdar stem, the imperfective stem, and the aorist stem. The suffix *-č* negates indicative forms. The prefix *t(A)-* negates non-indicative forms and is used by a limited number of strong verbs. The default structure to negate non-indicative forms is the periphrastic prefixal negative form which consists of a negated auxiliary and a periphrasis form. The past tense marker *-j/-ir* is farther away from the verbal stem than tense-aspectual markers. The past tense marker *-ir* which is used in a negative context follows the negative suffix *-č*. The participle marker *-j* expresses relative clauses, scopes over temporal-aspectual forms, and occurs farther away from the verbal stem.

3.3 A realization OT approach to affix order in Lezgian

In this section I present a realization OT account of the data and generalizations in section 2. The crucial constraint is the scope constraint which associates a semantic scopal relation with a linear order. Following Spencer 2003, I define the scope constraint as follows:

- (33) SCOPE: Given two scope-bearing features f_1 and f_2 , if f_1 scopes over f_2 , then I_2 which is the exponent of f_2 cannot be farther away from the same stem than I_1 which is the exponent of f_1 .

Within Paradigm Function Morphology (Stump 2001), Spencer 2003 defines the scope constraint as follows:

(34) The scope constraint (Spencer 2003: 643): Given a paradigm function evaluated for scope-bearing features f_i, f_j , if f_j scopes over f_i then $I_i > I_j$, where I_i, I_j are affix indexes associated respectively with f_i, f_j .

Notice that the difference between my definition and his is that I define the scope constraint in a negative way, i.e. an exponent in the scope of another exponent should not be farther away from the same stem than the exponent taking scope. By contrast, Spencer seems to define the scope constraint in a positive way, i.e. an exponent in the scope of another exponent should be closer to the stem ($I_i > I_j$).

I argue that it is better to define the scope constraint in a negative way. Remember the constraint *FEATURE FUSION which bans an exponent realizing more than one feature value. If the scope constraint were defined in a positive way, it would replace *FEATURE FUSION which should be an independent constraint. Consider (35). The example in (35) violates the scope constraint defined in a positive way because the affix realizing case is not closer to the noun than the affix realizing number if case scopes over number. It also violates *FEATURE FUSION because the affix in (35) realizes both case and number.



SCOPE	Yes
*FEATURE FUSION	Yes

If the scope constraint is defined in a negative way, the functions of the scope constraint and *FEATURE FUSION can be clearly separated. See (36). The example in (36) does not violate the scope constraint defined in a negative way because the affix realizing number is not farther away from the noun than the affix realizing case. It, however, violates *FEATURE FUSION.



SCOPE	No
*FEATURE FUSION	Yes

Consider the example in (37). The pronoun *him* realizes person, number, gender, and case features. It is not clear whether it violates the scope constraint defined in a positive way because all features are realized within a single root. The verb *like* raises the same question. It realizes tense and aspect features but it is not clear whether it violates the scope constraint defined in a positive way. The words *him* and *like* in (37) apparently do not violate the scope constraint defined negatively, i.e. an exponent in

the scope of another exponent should not be farther away from the same stem than the exponent taking scope. Both the words *him* and *like* violate *FEATURE FUSION because each of them realizes several features.

(37) *They like him.*

This general scope constraint may be further decomposed into more specific constraints such as SCOPE (case, number), SCOPE (tense, aspect), etc. For example, SCOPE (case, number) can be defined as follows:

(38) SCOPE (case, number): The exponent of number cannot be farther away from the same nominal stem than the exponent of case, for case scopes over number because case expresses the relation of an entity or a number of entities to other elements in the clause.

Baker 1985 proposes a mechanism called the “Mirror Principle” to associate syntactic operations with morphological structures. He discusses orders of passive and agreement affixes, orders of causative and reflexive-reciprocal affixes, orders of passive and applicative affixes with respect to orders of syntactic operations. See also Hyman 2003 for an OT approach to Bantu languages with the Mirror constraint.

It is hard to test the Mirror Principle in Lezgian based on the types of data discussed in Baker 1985. Lezgian has no passive or reflexive or reciprocal affixes. There is no agreement between adjective and noun or between verb and noun in Lezgian. Semantic scope-based ordering (Rice 2000) seems to be a more appropriate concept to use since the order of syntactic operations discussed by Baker or Hyman is not obvious in the Lezgian data. See Paster 2005 for similar reasoning.

3.3.1 Lezgian nominal inflection

Remember the generalization about affix order in Lezgian nominal inflection. All case markers are farther away from the nominal stem than number markers. The locatives scope over the localizations, which corresponds to the order in which the locative markers are farther away from the nominal stem than the localization markers. The generalization about the order of the locatives and localizations with respect to semantic scope generally holds in an earlier stage of the Lezgian language and remains visible in some of the case forms in the modern standard language. Consider the data in (39).

(39) Partial Lezgian nominal paradigm (*hül* ‘sea’)

	Plural
Absolutive	<i>hül-ér</i>
Ergative	<i>hül-ér-i</i>
Dative	<i>hül-ér-i-z</i>
Superrelative	<i>hül-ér-i-l-aj</i>
Superdirective	<i>hül-ér-a-l-di</i>

To account for the data in (39), I propose the constraint SCOPE (case, number). This constraint is a universal markedness constraint built on the generalization by Greenberg that “the expression of number almost always comes between the noun base and the expression of case” (Greenberg 1963: 112, cited in Bybee 1985: 34). We can interpret this constraint in the sense that case scopes over number because case expresses the relation of an entity or a number of entities to other elements in the clause.¹⁴ Additionally, I propose the following relevant constraints:

- (40) a. {absolutive}: -Ø: The absolutive case is realized by a zero suffix.
 b. {ergative, plural}: -i: The suffix -i realizes both the ergative and plural.¹⁵
 c. {dative}: -z: The dative case is realized by the suffix -z.
 d. {super}: -l: The super localization is realized by the suffix -l.
 e. {elative}: -aj: The elative is realized by the suffix -aj.
 f. {directive}: -di: The directive is realized by the suffix -di.
 g. {plural}: -Ar: The plural is realized by the suffix -Ar (either -ar or -er).
 h. SCOPE (locative, localization): The exponent of localization cannot be farther away from the same nominal stem than the exponent of locative because locative scopes over localization.

The ergative plural form *hül-ér-i* is derived through the constraints {ergative, plural}: -i, {plural}: -Ar, and SCOPE (case, number). See (41). Assume the input consists of the lexeme HÜL ‘sea’ with its nominal stem *hül-* and phonologically unrealized feature values pl(ural) and erg(ative). Candidate (b) **hül-i-er* is ruled out by SCOPE (case, number) because the plural marker -er is farther away from the nominal stem *hül-* than the case marker -i.

(41) *hül-ér-i*

HÜL ‘sea’, pl, erg <i>hül-</i>	{ergative, plural}: -i	{plural}: -Ar	SCOPE (case, number)
☞ a. HÜL ‘sea’, pl, erg <i>Hül -er -i</i>			
b. HÜL ‘sea’, pl, erg <i>hül -i -er</i>			*!

It is possible to rule out **hül-i-er* by a phonological constraint banning vowel hiatus (*VV). “In general, all non-initial syllables begin with exactly one consonant. However, there are a few Arabic loanwords like *düa* ‘prayer’, *šair* ‘poet’, in which there are vowel-initial medial syllables. These may be pronounced with a glottal stop (/dyʔa/, /ʃaʔir/, but the glottal stop may be omitted” (Haspelmath 1993: 41). I use *VV as a shorthand for the constraint banning vowel hiatus. See Rosenthal 1997 for a comprehensive discussion of constraints against prevocalic vowels.

The constraint *VV, however, cannot rule out some other candidates which violate the scope constraint. Consider the superdirective example *hül-ér-a-l-di* ‘onto the sea’ in an earlier stage of the Lezgian language. The super localization is realized by the suffix *-l* and the directive case is realized by the suffix *-di*. The phonological constraint *VV cannot rule out the ungrammatical candidate **hül-ér-a-di-l* which is phonologically well-formed, while the constraint SCOPE (locative, localization) can.

The dative form *hül-ér-i-z* and the superrelative form *hül-ér-i-l-aj* are built on the ergative plural form *hül-ér-i*. Consider the tableau in (42). Candidate (b) is ruled out by SCOPE (locative, localization) because the super exponent *-l* which is in the scope of the relative exponent *-aj* is farther away from the stem. Notice that an output-to-output correspondence constraint (Benua 1995, McCarthy and Prince 1995, Kenstowicz 1997, Kager 1999) may be involved to make the ergative plural form *hül-ér-i* copied by the superrelative plural form *hül-ér-i-l-aj*. See Chapter 5 for a relevant discussion.

(42) *hül-ér-i-l-aj* ‘from a position on the sea’

<i>hül-ér-i</i> , superrelative		{super}: <i>-l</i>	{relative}: <i>-aj</i>	SCOPE (locative, localization)
a.	super relative <i>hül-ér-i</i> <i>-l</i> <i>-aj</i>			
b.	super relative / \ <i>hül-ér-i</i> <i>-aj</i> <i>-l</i>			*!

To briefly summarize, the constraints SCOPE (case, number) and SCOPE (locative, localization) capture the generalizations about affix order in Lezgian nominal inflection.

3.3.2 Lezgian verbal inflection

Let us consider Lezgian verbal inflection. Recall the generalizations about affix order in Lezgian verbal inflection. The past tense marker *-j/-ir* is farther away from the verbal stem than tense-aspect markers. The past tense marker *-ir* which occurs in a negative context follows the negative suffix *-č*. The participle marker *-j* expresses relative clauses, scopes over temporal-aspectual forms, and occurs farther away from the verbal stem. The suffix *-č* negates indicative forms. The prefix *t(A)-* negates non-indicative forms and co-occurs with a limited number of strong verbs. The default structure to negate non-indicative forms is the periphrastic prefixal negative form which consists of a negated auxiliary and a periphrasis form. Consider (43).

(43) Partial Lezgian verbal paradigm (*fin* ‘go’)

	affirmative	negative	affirmative participle	negative participle
Masdar	<i>fi-n</i>	<i>te-fi-n</i>		
Infinitive	<i>fi-z</i>	<i>te-fi-z</i>		
Future	<i>fi-da</i>	<i>fi-da-č</i>	<i>fi-da-j</i>	<i>te-fi-da-j</i>
Past Future	<i>fi-da-j</i>	<i>fi-da-č-ir</i>		
Perfect	<i>fe-nwa</i>	<i>fe-nwa-č</i>	<i>fe-nwa-j</i>	<i>te-fe-nwa-j</i>
Past Perfect	<i>fe-nwa-j</i>	<i>fe-nwa-č-ir</i>		

To account for the partial paradigm in (43), I first propose the following constraints:

- (44) a. {masdar}: *-n*: The masdar is realized by the suffix *-n*.
 b. {infinitive}: *-z*: The infinitive is realized by the suffix *-z*.
 c. {future (habitual)}: *-da*: The habitual aspect is realized by the suffix *-da*.
 d. {perfect}: *-nwa*: The perfect is realized by the suffix *-nwa*.¹⁶
 e. {past/_[negative]}: *-ir*: The past tense in the negative environment is realized by the suffix *-ir*.
 f. {past}: *-j*: The past tense is realized by the suffix *-j*.
 g. {participle}: *-j*: The participle is realized by the suffix *-j*.
 h. {negative/_[+indicative]}: *-č*: Negation in the environment of indicative forms is realized by the suffix *-č*.
 i. {negative/_[FIN]}: *tA-*: Negation in the environment of the lexeme FIN is realized by the prefix *tA-*.
 j. SCOPE (tense, aspect): The exponent of aspect cannot be farther away from the same verbal stem than the exponent of tense because tense scopes over aspect.
 k. SCOPE (participle, tense & aspect): The exponents of tense and aspect cannot be farther away from the same verbal stem than the exponent of participle (relative clause marker) because participle (relative clause marker) scopes over tense and aspect.

Constraints (44a-i) are realization constraints that associate morphosyntactic or semantic feature values with phonological forms.

Constraints (44e, h, i) can be formulated in a more straightforward way without the notion of “environment”. For example, Constraint (h) can be reformulated in a way that the suffix *-č* realizes both negation and indicative. See Chapter 4 for a relevant discussion.

Constraints (44j-k) are scope constraints. The constraint SCOPE (tense, aspect) is a universal markedness constraint on the order of tense and aspect markers. Tense

scopes over aspect because “[t]ense places the situation [(habitual, progressive, perfect, etc.)] in time with respect to an established point in time, either the moment of speech, or some other point in time” (Bybee 1985: 28). In Bybee’s (1985) database, “[a]spect markers were found to be closer to the stem than tense markers in 8 languages, while the opposite order did not occur in the sample” (Bybee 1985: 34).

Notice that some Lezgian inflectional affixes may realize both tense and aspect. The aorist marker *-na*, for example, realizes both tense and aspect because the aorist is “the usual way to refer to perfective events in the past” (Haspelmath 1993: 142). The future exponent *-da* expresses either future or habitual situations and “[t]he Past is only compatible with the habitual meaning of the Future, not with the future meaning” (Haspelmath 1993: 142). The imperfective marker *-z(a)wa* may express progressive situations, habitual situations, and “states that obtain at the time of reference” (Haspelmath 1993: 140). In the past imperfective example Haspelmath 1993 gives, the imperfective refers to progressive situations. The order in which the past tense marker *-j/-ir* follows these tense-aspect markers satisfies SCOPE (tense, aspect) because the aspect exponent is not farther away from the verbal stem than the tense exponent.

The relative clause marker (e.g. *the one that* + clause) should scope over temporal-aspectual markers within a clause, which is reflected in the scope constraint SCOPE (participle, tense & aspect).

Let us see how the constraints in (44) account for the data in (43). Consider the tableau in (45). Candidate (45b) **fi-j-da* is ruled out by SCOPE (tense, aspect) because *-da* which expresses habitual situations in this case is farther away from the verbal stem *fi-* than the tense marker *-j*.

(45) *fi-da-j* (past habitual)

<i>fi-</i> , habitual, past	{future (habitual)}: <i>-da</i>	{past}: <i>-j</i>	SCOPE (tense, aspect)
☞ a. habitual, past <i>fi -da -j</i>			
b. habitual, past <i>fi -j -da</i>			*!

A similar analysis can be made of past perfect forms like *fe-nwa-j*, future participle forms like *fi-da-j*, and perfect participle forms like *fe-nwa-j* in (43) by using the relevant realization and scope constraints in (44).

The order of the negative suffix *-č* and tense or aspect markers is tricky. Bybee 1985 notices that “in some uses negation can resemble what we are calling mood, in that it can have the whole proposition in scope [(e.g. *it is not the case that* + clause)]” (p.176). See also the following Chichewa example which shows that negation scopes over other affixes.

(46) Multiple prefixation in Chichewa (Hyman 2003: 247)

Main root NEG- SUB- TNS- ASP- OBJ- stem ‘we will not just hit him’
 clause: *si-* *ti-* *dzá-* *ngo-* *mú-* *ményá* *-dzá-* ‘future’, *-ngo-* ‘just’

Bybee 1985 found that tense and aspect markers are universally closer to the stem than mood markers. We can first propose the following scope constraint and see how it works in Lezgian.

(47) SCOPE (negation, tense & aspect): The exponents of tense and aspect cannot be farther away from the same verbal stem than the exponent of negation because negation scopes over tense and aspect.

Let us use the negative past future (habitual) form *fí-da-č-ir* as an example for illustration. The form *fí-da-č-ir* apparently violates the constraint SCOPE (negation, tense & aspect) which requires the tense marker *-ir* not to be farther away from the same verbal stem than the negative marker *-č*. The negative marker *-č* precedes the past tense marker *-ir* because otherwise the ungrammatical form **fí-da-ir-č* would have a vowel hiatus which is generally banned in Lezgian. Consider the tableau in (48). Candidate (b) is ruled out by *VV though it satisfies the scope constraint.

(48) *fí-da-č-ir* (negative past habitual)

<i>fí-</i> , habitual, past, negative	neg/_[+ind]: <i>-č</i>	past/_[neg]: <i>-ir</i>	habit: <i>-da</i>	*VV	SCOPE
a. habitual, past, neg <i>fí</i> <i>-da</i> <i>-č</i> <i>-ir</i>					*
b. habitual, past, neg <i>fí</i> <i>-da</i> <i>-ir</i> <i>-č</i>				*!	

The ungrammatical form **fí-da-j-ir-č* which contains an epenthesized consonant [j] to avoid vowel hiatus can also be a candidate. It can be ruled out by the constraint DEP (C) which outranks SCOPE and bans consonant insertion.

A similar analysis based on the scope constraint can be made of negative future forms like *fí-da-č* and negative perfect forms like *fe-nwa-č*.

It is possible to formulate the constraint {past/_[negative]}: *-ir* simply as {past}: *-ir*, i.e. the past tense is realized by the suffix *-ir*. By doing that, we need to explain why *fí-da-č-ir* (negative past habitual) which violates SCOPE is better than **fí-da-j-č* which satisfies the scope constraint because the past tense suffix *-j* is closer to the verb stem than the negative suffix *-č*. One solution is to rank *COMPLEXCODA higher than SCOPE so that *COMPLEXCODA can rule out **fí-da-j-č*, although complex codas are common in Lezgian (e.g. /-jd/ *qejd* ‘remark’, /-wʃ/ *benewš* ‘violet’, /-fs/ *nefs* ‘thirst, desire’, /-χʃ/ *baxš* ‘dedication’). This might explain why the past tense marker

-j does not co-occur with the negative suffix *-č*. See the tableau in (49).

(49) *fi-da-č-ir* (negative past habitual)

<i>fi-</i> , habitual, past, negative	neg/_[+ind]: <i>-č</i>	past: <i>-ir</i>	past: <i>-j</i>	*COMPLEX CODA	SCOPE
a. habitual past neg X <i>fi</i> <i>-da</i> <i>-č</i> <i>-ir</i>					*
b. habitual past neg <i>fi</i> <i>-da</i> <i>-j</i> <i>-č</i>				*!	

The reason that the past tense marker *-ir* does not follow tense-aspect markers is clear (e.g. *fe-nwa-j* (past perfect), **fe-nwa-ir*). If the past tense suffix *-ir* follows a tense-aspect marker which always ends in a vowel, the consequent form (e.g. **fe-nwa-ir*) will lead to vowel hiatus which is generally banned in Lezgian.

Notice that the constraint {negative/_[+indicative]}: *-č* bears no subset relation to the constraint {negative/_FIN}: *tA-*, but the former needs to outrank the latter. Consider the tableau in (50). Candidate (b) **te-fe-nwa* is ruled out by {negative/_[+indicative]}: *-č*. Notice that the illicit form **te-fe-nwa-č* which bears two negation markers can be ruled out by *FEATURE SPLIT. Both {negative/_[+indicative]}: *-č* and {negative/_FIN}: *tA-* which introduce synthetic negation markers should outrank the less specific constraint introducing the periphrastic negative form which is the default negative structure in Lezgian.¹⁷

(50) *fe-nwa-č* (negative perfect)

<i>fe-</i> , perfect, negative	Negative/_[+indicative]: <i>-č</i>	Negative/_FIN: <i>tA-</i>
a. perfect, negative <i>fe</i> <i>-nwa</i> <i>-č</i>		*
b. negative perfect <i>te-</i> <i>fe</i> <i>-nwa</i>	*!	

There is, however, a problem with the constraint SCOPE (negation, tense & aspect). Bybee 1985 does not consider negation a member of the mood category in general. She says that “[a] major difference between the conceptual category of negation and other moods is that the scope of negation is highly variable, and in some cases it can have a single lexical item in its scope [(e.g. *dis-like*)]” (p.176).

In Turkish, for example, the negation marker may be closer to the verb stem than tense markers (51).

- (51) *İstanbul-a git-me-yecek-ti-m.*
 Istanbul-DAT go-NEG-FUT-PAST-1SG
 ‘I wasn’t going to go to Istanbul.’ (Wilson and Saygin 2001)

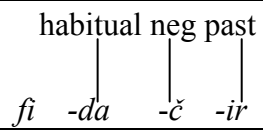
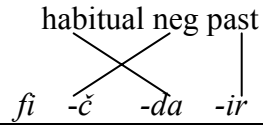
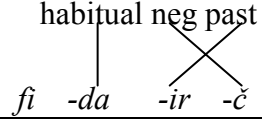
Another reason for the difficulty in determining the scope of negation with respect to tense and aspect is that in many languages negative markers are prefixes (see Bybee 1985 for a relevant discussion). It is therefore difficult to tell whether a negative prefix is farther away from the stem than tense and aspect suffixes.

If there does not exist a universal scope constraint like SCOPE (negation, tense & aspect), we can only stipulate the position of a negation marker with respect to tense and aspect markers. In Lezgian, there does not seem to be language-particular evidence that negation markers are farther away from the verb stem than tense markers. We may refer to language-particular constraints such as those in (52).

- (52) a. Negation > Tense: The exponent of negation cannot be farther away from the same stem than the exponent of tense.
 b. Aspect > Negation: The exponent of aspect cannot be farther away from the same stem than the exponent of negation.

The two constraints in (52) will put the negative marker *-č* between a (tense-)aspectual marker and a tense marker. The constraint Aspect > Negation needs to outrank Negation > Tense because in Lezgian, tense-aspect markers such as *-na* {aorist} which realizes both tense and aspect precede the negative marker *-č* (*fe-na-č*, **fe-č-na*). See the tableau in (53). Notice that compared to *fi-da-č-ir*, another illicit candidate **fi-da-č-j* which has the past tense marker *-j* violates not only *COMPLEXCODA but also the universal Sonority Sequencing Principle (Kenstowicz 1994) which requires a coda cluster to have a contour with falling sonority. The coda cluster *[-čj] is illicit in Lezgian.

(53) *fi-da-č-ir* (negative past habitual)

<i>fi</i> -, habitual, negative, past	habit: <i>-da</i>	negative/_ [+indicative]: <i>-č</i>	past: <i>-ir</i>	Asp > Neg	Neg > Tense
a. 					
b. 				*!	
c. 					*!

The two constraints in (52) which put the negation marker between tense and aspect markers are similar to the constraint *TEMPLATE* in Hyman 2003 which stipulates affix order. Hyman shows that the order of several types of affixes in Bantu languages strictly conforms to a fixed morphological template, i.e. Causative suffix - Applicative suffix - Reciprocal suffix - Passive suffix (the order from left to right corresponds to the distance from the closest to the farthest between a stem and a suffix).¹⁸ Hyman’s OT model is quite similar to our model in that the input consists of a stem and phonologically unrealized morphosyntactic feature values which are spelt out in the output. Hyman, however, does not talk about realization constraints which associate abstract morphosyntactic feature values with phonological forms. The constraint *TEMPLATE* is, for example, defined as follows: “A morphosyntactic input {CAUS[ATIVE], APP[LICATIVE]} is realized according to [the morphological template] *CARP*, i.e. *-its-il-*” (Hyman 2003: 249). As we can see, the constraint *TEMPLATE* has two functions. One is to stipulate affix order just like the constraints in (52)¹⁹ and the other is to spell out abstract morphosyntactic feature values. The second function of *TEMPLATE* can be fulfilled with more specific realization constraints (e.g. Causative: *-its*) in our model.

A similar analysis based on language-particular constraints on affix order can be made of the negative future form *fi-da-č* and the negative perfect form *fe-nwa-č* in the paradigm in (43).

If we need to use templatic constraints to stipulate the position of the negative suffix *-č*, orders of affixes in Lezgian can be analyzed as a “mixed Scope-Template system” (cf. Paster 2005).

To briefly summarize, we can describe the patterns about affix order in Lezgian verbal inflection by using both universal scopal and phonological constraints and language-particular realization and templatic constraints.

3.4 Other approaches

In this section, I compare our realization OT approach with several alternative approaches to affix order and argue for its advantages.

In Paradigm Function Morphology (PFM) (Stump 2001), Stump places affixes in linear orders by using rule blocks. An affix which is closer to a stem is placed in a rule block which precedes another rule block containing an affix which is farther away from the same stem. This approach is essentially similar to that in Anderson (1992) which cyclically implements affixation. See also Stump 1993 which places affixes in position classes (each of which represents a rule block). Consider a Finnish example in (54) which is described within the framework of PFM. “The Finnish noun form *talo-i-ssa-ni* ‘in my houses’ consists of the root *talo*, followed by Plural, Inessive Case and 1Sg Possessor affixes” (Spencer 2003: 630). The output from Rule Block I (*X_i*) becomes the input to Rule Block II (*X*).

- (54) *talo-i-ssa-ni* ‘in my houses’ (Finnish)
 (a) Block I, [NUM: Pl], [CLASS: Noun] (X) = *Xi*
 (b) Block II, [CASE: Iness], [CLASS: Noun] (X) = *Xssa*
 (c) Block III, [POSSESSOR: 1sg], [CLASS: Noun] (X) = *Xni*

One of the shortcomings of this approach is that by simply labeling each rule block with a number from small to big we miss universal generalizations with respect to scope, e.g. case scopes over number, so a case exponent should not be closer to the same stem than a number exponent. Spencer 2003 also notices this problem and proposes the general scope condition imposed on rule blocks. See (34) (repeated in (55)).

- (55) The scope constraint (Spencer 2003: 643): Given a paradigm function evaluated for scope-bearing features f_i, f_j , if f_j scopes over f_i then $I_i > I_j$, where I_i, I_j are affix indexes associated respectively with f_i, f_j .

The scope constraint proposed by Spencer is a redundant strategy for PFM given that the order of rule blocks is determined by the distance between a stem and an affix on a language-particular basis. PFM might consider the scope constraint a sole factor in determining the order of rule blocks to avoid analytical redundancy. But it is not clear how the scope constraint interacts with a templatic constraint in PFM which is encoded in the order of rule blocks, which should override the scope constraint. It is therefore not clear how PFM solves this paradoxical situation: on the one hand, the scope constraint determines the order of rule blocks while it needs to be overridden by the order of rule blocks on the other hand. By contrast, the scope constraint points toward our realization OT approach which not only eliminates cyclic affixation but also encodes language universals in constraint rankings.

It is also possible to account for affix order in Lezgian via syntactic approaches. For example, in Distributed Morphology (DM) (Halle and Marantz 1993, Embick and Noyer 2001) which revels in every case of syntax-morphology interpenetration,²⁰ the scope generalization can be automatically assumed in syntactic structures with elements which take scope c-commanding elements within the scope.²¹ Under a DM approach, a morphological structure is derived from a syntactic structure that may undergo processes such as lowering of morphosyntactic elements in a syntactic tree or head-to-head movement followed by Vocabulary Insertion which realizes abstract morphosyntactic feature values and places affixes in linear orders. Compared to this approach, the proposed OT realization approach in accounting for the Lezgian inflectional system not only spells out universal generalizations about scope effects and universal restrictions on phonotactics, but also avoids cyclic derivation of a morphological structure although it remains controversial whether cyclicity is necessary in our grammatical architecture.

The realization OT approach proposed in this chapter captures scope generalizations more readily than a framework based on the generation of output morphosyntactic feature values (Grimshaw 1997b, 2001, Wunderlich 2001). This kind

of framework lacks a systematic mechanism to spell out these abstract feature values within the framework of OT. Without introducing extra and distinct machinery such as Vocabulary Insertion which is used in a cyclic framework like Distributed Morphology, it is difficult to incorporate into the framework the scope constraint which associates semantic scope with the order of phonological forms. By contrast, the proposed realization OT approach naturally incorporates the scope constraint because this framework is built on a mechanism which associates abstract morphosyntactic or semantic feature values with phonological forms.

In a morphologically restricted OT model (McCarthy and Prince 1993a, b, McCarthy and Prince 1995), it is considered that morphological information such as “affix”, “root”, and “stem”, and phonological information related to phonetic/phonological features, segments, and suprasegmental properties constitute enough input and output information for the grammar to process. It is thus hard to see how to incorporate into this model the scope constraint which crucially relies on morphosyntactic feature values if we do not introduce morphosyntactic feature values into both the input and output.²² Within a morphologically restricted OT model, Paster 2005 uses the scope constraint to account for affix order in Pulaar, a West Atlantic language spoken across a wide area of West Africa. She has to assume that morphosyntactic feature values are present in both the input and output in order for the scope constraint to work, though we do not see such feature values in her tableaux. There seems to be no big difference in accounting for affix order by means of the scope constraint between a morphologically restricted OT approach which introduces morphosyntactic feature values and the proposed realization OT approach. Putting aside trivial issues such as what phonological representations a morphologically restricted OT model should posit in the input, we should notice that phonological information is introduced through the input in a morphologically restricted OT model but through realization constraints in our model. As we have argued in the previous chapter, a realization OT approach more readily accounts for cases of OCP-triggered (allo)morph selection which involve the emergence of a morphosyntactically less specific form while such a selection needs to be stipulated in the input under a morphologically restricted OT approach. We will continue to argue in the following chapter that a morphologically restricted OT approach fails to account for blocking and extended morphological exponence both of which can be readily explained under a unified approach within our realization OT model.

3.5 Conclusion

In this chapter I study Lezgian inflectional morphology with a focus on affix order. It is found that case markers are outside number markers. Locative markers which scope over localization markers are farther away from the nominal stem. The past tense marker is outside tense-aspect markers. Participles which express relative clauses are outside temporal-aspectual affixes. The negative marker in the indicative environment occurs between past-tense and tense-aspect markers. Additionally, the past tense suffix *-ir* does not appear in an affirmative context or follow tense-aspect markers, which always end in a vowel. I show that these generalizations can be

captured by both universal scopal and phonological constraints and language-particular realization and templatic constraints. If the position of the negative suffix *-č* reflects a templatic account, affix order in Lezgian can be analyzed as a “mixed Scope-Template system” (cf. Paster 2005).

I compare our realization OT model with the above-mentioned constraint types to other approaches. I show that this approach not only captures universal scope and phonological generalizations but also avoids cyclic derivation of morphological structures, although it remains controversial whether cyclicity is necessary in our grammatical architecture.

There are remaining issues. For example, based on Julien 2000, Trommer 2003 argues that although Bybee’s 1985 generalization that tense markers are not inside aspect markers is universally true, it seems that the order Aspect-Verb-Tense is “virtually non-existent.” He thus argues for a cyclic syntactic approach based on head-to-head movement which posits Tense-Aspect-Verb as an underlying structure. He shows that Aspect-Verb-Tense is not a possible stage in such a syntactic derivation. If that is true, it shows that in addition to the scope constraint, additional constraints should be imposed on the order of, for example, tense and aspect markers if we want to maintain a restrictive non-cyclic OT account.

Notes

¹ The suffix *-a* is “[a]lso used in a small number of common nouns (e.g. *apaj*, obl. *apaj-a* ‘father-in-law’)” (Haspelmath 1993: 75).

² “[T]here are a few non-derived nouns that take the oblique suffix *-i* in the singular. Some of them end in *-(u)n* or *-r*, which suggests that they represent old lexicalized Masdar forms and former *pluralia tantum* which have been reanalyzed as singulars” (Haspelmath 1993: 76). See (i).

(i) <i>dugun</i>	obl.	<i>dugún-i</i>	‘valley’
<i>ğucar</i>	obl.	<i>ğucár-i</i>	‘god’
<i>q’ular</i>	obl.	<i>q’ulár-i</i>	‘mousetrap’
<i>cwal</i>	obl.	<i>cwál-i</i>	‘seam’

“In some monosyllabic nouns, the oblique suffix *-i* is stressed” (Haspelmath 1993: 76). See (ii).

(ii) <i>k’wač</i>	obl.	<i>k’wač-i</i>	‘foot’
<i>k’wal</i>	obl.	<i>k’wal-i</i>	‘house’
<i>k’an</i>	obl.	<i>k’an-i</i>	‘bottom’
<i>kal</i>	obl.	<i>kal-i</i>	‘cow’
<i>kac</i>	obl.	<i>kac-i</i>	‘cat’

³ “When the plural suffix is not stressed, its vowel is syncopated before a following vowel-initial oblique stems suffix” (Haspelmath 1993: 73).

⁴ “In a few isolated cases *-rA* is also used with nouns that denote people and inanimate objects” (Haspelmath 1993: 76). See (i).

(i) <i>luk'</i>	obl.	<i>luk'-ra</i>	‘slave’
<i>čam</i>	obl.	<i>čam-ra</i>	‘bridegroom’
<i>warz</i>	obl.	<i>wac-ra</i>	‘moon; month’

⁵ The initial consonant of each variant *-c'i/-c'í/-č'i/-č'í/-ž'i* seems to agree with that of a preceding root. See (i) (from Haspelmath 1993: 77).

(i) <i>par</i>	obl.	<i>par-cí</i>	‘load’
<i>k'ar</i>	obl.	<i>k'ar-c'í</i>	‘rolling pin’
<i>čar</i>	obl.	<i>čar-č'í</i>	‘paper; letter’
<i>č'ar</i>	obl.	<i>č'ar-č'í</i>	‘hair’
<i>žin</i>	obl.	<i>žin-ž'í</i>	‘ghost’

⁶ The oblique suffix *-i* after the abstract suffix *-wil* is lowered to *-e*.

⁷ The symbol \tilde{q} stands for [q], an unaspirated and non-labialized uvular.

⁸ “Many monosyllabic loanwords take the default suffix *-ar* (e.g. *pek*, *pék-ar* ‘silk’). A few native monosyllabic nouns also show this suffix (e.g. *ner*, *nér-ar* ‘noses’) ... A few monosyllabic nouns with a back stem vowel form their plural idiosyncratically in *-ér* (e.g. *kal-ér* ‘cows’). Nouns derived with the suffix *-wal* form their plural in *-er*. (e.g. *šád-wal*, *šád-wil-er* ‘lit. joys’)” (Haspelmath 1993: 71-72).

Lezgian has other plural markers. “A few monosyllabic nouns have the suffix *-lAr*. (e.g. *bağ*, *bağ-lár* ‘gardens’; *el*, *el-lér* ‘nations’) ... A few nouns have the suffix *-Arar* (e.g. *p'uz*, *p'uz-árar* ‘lips’; *žin*, *žin-érar* ‘ghosts’) ... The substantivizing affix on adjectives is *-bur* in the plural (e.g. *jaru* ‘red’, *jaru-di* ‘red one’, *jaru-bur* ‘red ones’)” (Haspelmath 1993: 72-73). Plurals of certain Russian loanwords may end in *-ajar*, *-ijar*, or *-(ja)r*.

⁹ Inflectional suffixes such as *-zwa* and *-nwa* may be derived from *-zawa* and *-nawa* respectively through “post-tonic vowel syncope”, i.e. “/a/ drops in a medial post-tonic syllable if preceded by no more than one consonant (*raxázwa*), otherwise it is preserved (*šéx̂nawa*)” (Haspelmath 1993: 39). Haspelmath 1993 also points out that the precise phonological conditions for post-tonic vowel syncope are very hard to identify and this syncope “seems to be restricted to inflectional suffixes ... in some varieties of the language, the /a/ has been syncopated [even after an unstressed high vowel] (e.g. *šéx̂-izwa*)” (Haspelmath 1993: 39).

¹⁰ Though the thematic vowel of a masdar stem is always high, it is difficult to predict the specific vowel. E.g. The verbal root *f-* takes the thematic vowel [i] in its masdar stem *f-í-*.

¹¹ The aorist participle is realized by “-*aj* for weak verbs (or rarely and archaically -*ur*) and -*r/-j/-ji* for strong verbs. The form -*r* is used after a high thematic vowel (u, ü), and -*j/-ji* is used after a low thematic vowel (a, e). The form -*j* is used after polysyllabic stems, -*ji* is used after monosyllabic stems” (Haspelmath 1993: 131).

¹² The verb *t’ün* ‘eat’ combines with the negative marker *tü-* (i.e. *tü-t’ün*).

¹³ “[The future] used to be a very general non-past form. After the Imperfective took over first the progressive and later the general present meanings, all that was left for this form was the future and habitual meanings” (Haspelmath 1993: 130).

¹⁴ See Bybee 1985 for a different interpretation with respect to the order of case and number. Bybee 1985 hypothesizes that “[a] meaning element is *relevant* to another meaning element *if the semantic content of the first directly affects or modifies the semantic content of the second*” (p.13). She further remarks that “the expression of number occurs closer to the noun base because it is more relevant to the meaning of the noun. Number has a direct effect on the entity or entities referred to by the noun. Case, on the other hand, has no effect on what entity is being referred to, but rather only changes the relation of that same entity to the other elements in the clause” (p.34).

¹⁵ The Lezgian nominal inflection schema, noun + plural + ergative case can be found in other languages. I am very grateful to Alice C. Harris who provides me with the following data.

(i) Tabasaran: *š:aw* ‘nail’ (Magometov 1965: 112-113)

	Singular	Plural
Absolutive	<i>š:aw</i>	<i>š:aw-ar</i>
Ergative	<i>š:aw-di</i>	<i>š:aw-ar-i</i>
Genitive	<i>š:aw-di-n</i>	<i>š:aw-ar-i-n</i>
Dative	<i>š:aw-di-s</i>	<i>š:aw-ar-i-s</i>

(ii) Aghul: *k’ar* ‘palka, drova’ (Magometov 1970: 73)

	Singular	Plural
Absolutive	<i>k’ar</i>	<i>k’ur-ar</i>
Ergative	<i>k’ar-u</i>	<i>k’ur-ar-i</i>
Genitive	<i>k’ar-u-n</i>	<i>k’ur-ar-i-n</i>
Dative	<i>k’ar-u-s</i>	<i>k’ur-ar-i-s</i>

It is also possible to treat the ergative plural marker -*i* as a meaningless oblique suffix. I leave open the question whether it is better to treat the suffix -*i* as an ergative case marker or meaningless oblique suffix.

¹⁶ I put aside the discussion of post-tonic vowel syncope, through which inflectional suffixes such as -*nwa* may be derived from inflectional suffixes such as -*nawa*, since our focus is on affix order instead of phonological alternation.

¹⁷ See Kiparsky 2005 for a review of various approaches to the blocking of periphrastic structures by synthetic forms.

¹⁸ Hyman’s suggestion for the existence of such a morphological template is that:

[T]he ideal alignment of [order of affixation with semantic compositionality, syntactic mirror principle, morphological layering, and phonological cyclicity] is just too much to ask of actual speaker-hearers faced with having to pack and unpack morphologically complex forms in real time. It is not possible, and it is often not even desirable for the order of affixation to be taken so literally. In fact, much of the scope relations is either trivially predictable from the lexical semantics or the discourse context or non-consequential, even indeterminate. (Hyman 2003: 263)

¹⁹ See also Paster 2005 which argues that the constraint *TEMPLATE* can be decomposed into specific templatic constraints.

²⁰ The mechanism to derive a morphological structure in Distributed Morphology embodies the claim in Baker 1985, i.e. morphology and syntax interact in one component, i.e. the *syntax proper*.

Grimshaw 1986 argues that the Mirror Principle does not necessarily show that morphology and syntax interact in one grammatical component. Instead, she suggests that morphology and syntax belong to two different grammatical components.

Hyman 2003 argues that at least some affix-orderings in Bantu should be explained in the *morphology proper* and concludes that “the Mirror Principle may not be universal in the ‘no exceptions’ sense, but rather in the (violable) OT sense” (p.260).

²¹ Rice 2000 argues that morpheme order in the Athapaskan verb is the result of semantic scope, where scope is reflected in linear order. She relates semantic scope to syntactic c-commanding. See Spencer 2003 for criticisms of Rice’s approach to Athapaskan verbal morphology.

²² In an interesting paper within the framework of a morphologically restricted OT model, Russell remarks that “[p]erhaps Bybee’s 1985 generalizations can be expressed as a set of Generalized Alignment constraints that form part of Universal Grammar, which may sometimes be violated when forced by some higher ranked constraint” (Russell 1997: 130). However, it is hard to see how alignment constraints (McCarthy and Prince 1993a) can account for the scope generalization defined in a negative way, i.e. an exponent in the scope of another exponent cannot be farther away from the same stem.

Chapter Four

A realization OT approach to blocking and extended morphological exponence

4.1 Introduction

Blocking and extended morphological exponence have been widely discussed in the literature on inflectional morphology. Blocking (Anderson 1986, Noyer 1992, Stump 2001, among many others) refers to a phenomenon in which a rule or affix *bleeds* the application of another rule or affix that expresses a similar (or the same) morphosyntactic feature value set to that expressed by the bleeding rule or affix. Extended (morphological) exponence (Matthews 1991, Noyer 1992, Anderson 2001, Stump 2001, among many others), on the other hand, refers to a phenomenon in which a morphosyntactic or semantic feature value is realized by more than one exponent in the same word.

Within realization models, two major approaches to blocking and extended exponence have been proposed. Noyer (1992, 1997) proposes a mechanism called “feature discharge” to account for some cases of blocking in inflectional morphology. According to this theory, once a morphosyntactic feature value is discharged or spelled out by an affix, it will be no longer available for further realization and therefore blocks the insertion of an affix that realizes the same morphosyntactic feature value. In order to allow extended exponence, Noyer makes a further distinction between primary and secondary exponents. That is, an affix which realizes a morphosyntactic feature value as a secondary exponent presumes another affix which realizes the same morphosyntactic feature value as a primary exponent. Extended exponence in Noyer’s framework therefore involves occurrences of both a primary and secondary exponent.

By contrast, Stump 2001 accounts for both blocking and extended exponence within a paradigm-based model which consists of realization rules that associate morphosyntactic feature values with phonological forms. In this theory, an affixal exponent is introduced by a realization rule to fill in a templatic slot. Realization rules which are supposed to fill in the same slot are placed in one rule block. Blocking is assumed to take place among realization rules that compete for the same morphotactic slot. In other words, blocking is assumed to take place within a single rule block. Additionally, Stump proposes the Pāṇinian Determinism Hypothesis, i.e. competition among realization rules within a single rule block can only be determined by Pāṇini’s Principle which requires a realization rule to preempt others if it applies to a more specific morphosyntactic feature value set.¹ In this framework, some cases of extended exponence involve more than one rule block. That is, exponents among

whose morphosyntactic or semantic feature value sets there exists a subset relation are placed in different rule blocks.

Notice that in the above-mentioned approaches both of which admit the validity of Pāṇini's Principle, distinct machinery needs to be introduced in order to allow extended exponence. Noyer (1992, 1997) resorts to a distinction between primary and secondary exponents while Stump (2001) resorts to multiple rule blocks.

In this chapter, I argue for our realization OT approach and show that it provides a unified account of both blocking and extended exponence without recourse to either a distinction between primary and secondary exponents or multiple rule blocks. The key device is the markedness constraint *FEATURE SPLIT which bans the realization of a morphosyntactic or semantic feature value by more than one exponent. The ranking of *FEATURE SPLIT and realization constraints which express the same morphosyntactic or semantic feature value(s) determines whether we should observe blocking or extended exponence. If *FEATURE SPLIT ranks lower than realization constraints which express the same feature value(s), then we should expect extended exponence. Otherwise, we should expect blocking among affixes.

The organization of this chapter is set as follows. In section 4.2, I discuss both the markedness constraint *FEATURE SPLIT and Pāṇini's Principle. I show that *FEATURE SPLIT is a more general constraint banning the realization of a morphosyntactic or semantic feature value by more than one exponent while Pāṇini's Principle more specifically applies to exponents among whose morphosyntactic or semantic properties there exists a subset relation. Additionally, I show that *FEATURE SPLIT *cannot* be replaced by alignment constraints (McCarthy and Prince 1993, Russell 1997, Grimshaw 2001) to account for blocking. In section 4.3, I compare our realization OT approach with the approaches to blocking and extended exponence in Noyer (1992, 1997) and Stump (2001) and show that the proposed realization OT approach not only readily captures both blocking and extended exponence by a unified device but also avoids extra machinery such as a distinction between primary and secondary exponents and multiple rule blocks. I discuss other alternative morphological approaches to blocking and extended exponence in section 4.4 and conclude in section 4.5.

4.2 *FEATURE SPLIT and Pāṇini's Principle

In inflectional morphology, Pāṇini's Principle is commonly introduced to account for cases in which an exponent spelling out a more specific set of morphosyntactic or semantic feature values blocks the occurrence of another exponent spelling out a subset of these feature values. Pāṇini's Principle was originally brought to modern attention in Kiparsky 1973 and is now known under various names such as the subset principle, the proper inclusion principle, the elsewhere condition, etc. Pāṇini's Principle can also be interpreted in different ways (see Anderson 1986). For example, an irregularly inflected form may block a regular realization rule which applies to a superset of stems including that of the irregularly inflected form.

In English, for example, the plural form of the noun OX is the irregular form *ox-en*. The word *oxen* not only expresses the plural feature value but also lexical

properties of the noun OX. It blocks the regular plural suffixation in English, i.e. the *s*-suffixation which is expected to apply to all English nouns including OX. The word *ox-en* thus blocks ungrammatical words like **ox-es* and **ox-en-s* which contain the suffix *-s*.

This case of blocking in English can be captured within Optimality Theory if we convert part of Pāṇini’s Principle into a markedness or well-formedness constraint. We can call this constraint “Subset Blocking” and formulate it as follows:

- (1) SUBSET BLOCKING: An exponent (Exponent₁) cannot co-occur with another exponent (Exponent₂) if the latter (Exponent₂) realizes a subset of feature values that are realized by the former (Exponent₁).

Additionally, I propose the following constraints:

- (2) a. {pl}: *-s*: The plural feature value is realized by the suffix *-s*.
 b. LEXICON: Information stored in the lexicon should be spelled out in the output.

The constraint LEXICON is an “output-to-output correspondence constraint involving lexical conservatism” (Steriade 2005). In this case, it equals saying that the plural form of the noun OX which is stored in the lexicon should be spelled out in the output. Because this constraint introduces a form (*oxen*) which not only realizes the plural feature value but also expresses lexical properties of the noun OX, it should outrank the constraint {pl}: *-s* in the spirit that a form or rule associated with a more specific feature value set has priority to apply.

Consider the following tableau. Candidate (a) wins although it violates the lower ranked constraint {pl}: *-s* because *-s* is not in the output. Candidate (b) is ruled out by SUBSET BLOCKING because the exponent *oxen* which expresses not only the plural feature value but also lexical properties of the lexeme OX co-occurs with the suffix *-s* which realizes a subset of feature values that are realized by *oxen*. Candidate (c) violates the constraint LEXICON because *oxen* which is stored in the lexicon is not spelled out in the output.

(3) *oxen*

Input: OX, pl <i>ox</i> Lexicon: OX, pl: <i>oxen</i>	LEXICON	SUBSET BLOCKING	{pl}: <i>-s</i>
☞ a. $\begin{array}{c} \text{OX} \quad \text{pl} \\ \diagdown \quad \diagup \\ \text{oxen} \end{array}$			*
b. $\begin{array}{c} \text{OX} \quad \text{pl} \\ \diagdown \quad \diagup \\ \text{oxen} \quad \text{-s} \end{array}$		*!	
c. $\begin{array}{c} \text{OX} \quad \text{pl} \\ \downarrow \quad \downarrow \\ \text{ox} \quad \text{-s} \end{array}$	*!		

For linguists who do not believe in the existence of the lexicon, the constraint LEXICON may be replaced by a realization constraint in (4) which is more specific than {pl}: -s.

- (4) {pl / OX}: -en: The plural feature value in the environment of the lexeme OX is realized by the suffix -en.

The constraint in (4) can also be formalized in such a way that the suffix -en realizes both the plural feature value and a diacritic feature (<en-class>) associated with nouns like OX. Such reformulation won't affect our results. I put aside the issue as to which approach, either the one in (3) or the one in (5), is better since it would otherwise invoke a comparison between models based on the storage of forms and models completely based on the generation of new forms, which is beyond the scope of this chapter.

The modified grammar will also choose the correct output *oxen* (5). Candidate (b) violates the constraint SUBSET BLOCKING because the suffix -en co-occurs with the suffix -s which realizes the plural feature value, which is subsumed in the feature value set realized by the suffix -en whether -en realizes a diacritic feature or not.

(5) *oxen* (no lexicon)

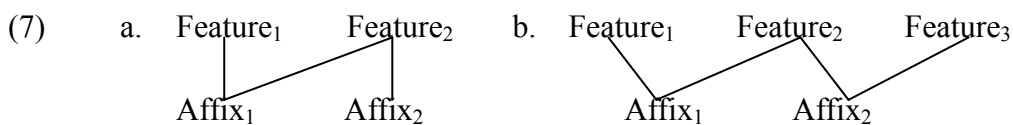
Input: OX, pl <i>ox</i>	{pl / OX}: -en	SUBSET BLOCKING	{pl}: -s
☞ a. OX pl <i>ox -en</i>			*
b. OX pl <i>ox -en -s</i>		*!	
c. OX pl <i>ox -s</i>	*!		

Notice that *FEATURE SPLIT can replace SUBSET BLOCKING to account for cases in which *oxen* blocks occurrences of **oxes* and **oxens*. See (6). Candidate (b) is ruled out by *FEATURE SPLIT because the plural feature value is realized by both the word *oxen* and the suffix -s.

(6) *oxen*

Input: OX, pl <i>ox</i> Lexicon: OX, pl: <i>oxen</i>	LEXICON	*FEATURE SPLIT	{pl}: -s
a. OX pl \ / <i>oxen</i>			*
b. OX pl / <i>oxen</i> -s		*!	
c. OX pl <i>ox</i> -s	*!		

The constraints SUBSET BLOCKING (SB) and *FEATURE SPLIT (*FS) are distinct but related mechanisms. There is a subset relation between SUBSET BLOCKING and *FEATURE SPLIT. That is, SUBSET BLOCKING bans a subset of forms that are banned by *FEATURE SPLIT. Assume that in (7) Features (1-3) are distinct. (7a) violates SUBSET BLOCKING because Affix₁ which realizes both Feature₁ and Feature₂ co-occurs with Affix₂ which only realizes Feature₂, which is subsumed in the feature value set realized by Affix₁. (7a) also violates *FEATURE SPLIT because Feature₂ is realized by both Affix₁ and Affix₂. By contrast, (7b) does not violate SUBSET BLOCKING because there is no subset relation between the feature value sets realized by Affix₁ and Affix₂. (7b), however, violates *FEATURE SPLIT because Feature₂ is realized by both Affix₁ and Affix₂. (7b) is a case of *overlapping exponence* (Matthews 1991) in which Feature₂ is shared by both Affix₁ and Affix₂ which express different morphosyntactic feature value sets.



SB	yes	no
*FS	yes	yes

A question arises as to whether markedness constraints like *FEATURE SPLIT can be replaced by alignment constraints (McCarthy and Prince 1993, Russell 1997, Grimshaw 2001) to account for morphological blocking. For example, in the case of *oxen* blocking **oxens*, we might put the suffixes *-en* and *-s* in a set $\{-en > -s\}$ in which *-en* has priority to be spelled out. We can modify Bonet's theory by assuming that it is possible for both *-en* and *-s* to appear in the output. The illicit form **oxens* can be ruled out by an alignment constraint which requires the plural marker to occur adjacent to the noun since *-s* is not adjacent to the noun (cf. Russell 1997, Grimshaw

2001). The relevant constraints are shown as follows.

- (8) a. PRIORITY: A form on the left of “>” has priority over a form on the right in being spelled out. (Bonet 2004)²
 b. N-Pl: A plural marker should occur adjacent to the noun. (cf. Russell 1997, Grimshaw 2001)

Consider the following tableau. Candidate (b) is ruled out because *-s* is not adjacent to the noun *ox*. Candidate (c) is ruled out because the suffix *-en* which has priority to be spelled out is not in the output. In this case, it seems that an approach based on alignment constraints can account for the blocking of **oxens* by *oxen*.

(9) *oxen*

<i>ox</i> + {-en > -s}	PRIORITY	N-Pl
☞ a. <i>ox-en</i>		
b. <i>ox-en-s</i>		*!
c. <i>ox-s</i>	*!	

I show that *FEATURE SPLIT cannot be replaced by alignment constraints. Notice that in the case of *oxen* blocking **oxens* it may be assumed that the suffixes *-en* and *-s* compete for one affixal slot, i.e. the position right after the noun *ox*. Alignment constraints therefore seem to be able to describe such a competition.

The constraint *FEATURE SPLIT, however, works better than alignment constraints in explaining cases in which both blocking and blocked exponents do not compete for one slot. Consider the Lezgian data in (10). The suffix *-č* realizes both negation and indicative mood. The prefixal negation marker *t-* occurs elsewhere. Notice that in the environment of the verb AWUN, the two negation markers *t-* and *-č* are in complementary distribution.³

(10) *awun* ‘do’ (adapted from Haspelmath 1993: 135)

	affirmative	negative
Masdar	<i>awú-n</i>	<i>t-awú-n</i>
Optative	<i>awú-raj</i>	<i>t-awú-raj</i>
Infinitive	<i>ijí-z</i>	<i>t-ijí-z</i>
Imperfective	<i>ijí-zwa</i>	<i>ijí-zwa-č</i>
Imperfective Participle	<i>ijí-zwa-j</i>	<i>t-ijí-zwa-j</i>
Future	<i>ijí-da</i>	<i>ijí-da-č</i>
Hortative	<i>ijí-n</i>	<i>t-ijí-n</i>
Aorist	<i>awú-na</i>	<i>awú-na-č</i>
Perfect	<i>awú-nwa</i>	<i>awú-nwa-č</i>
Aorist Participle	<i>awú-r</i>	<i>t-awú-r</i>
Aorist converb	<i>awú-na</i>	<i>t-awú-na</i>

It is hard to see how an alignment constraint rules out illicit forms like **t-iji-zwa-č* which contain both prefixal and suffixal negation markers because *t-* and *-č* do not compete for one position. By contrast, *FEATURE SPLIT readily rules out **t-iji-zwa-č* since negation is realized by both *t-* and *-č*. (See also Noyer (1992, 1997) for cases of *discontinuous bleeding* in which an affix is argued to block another affix even though they belong to different positions.)

Additionally, compared to various types of alignment constraints such as N-PLURAL, PERSON RIGHT (a person marker should be at the rightmost edge), etc., *FEATURE SPLIT, which is a universal mechanism underlying every language, is formulated in a more consistent and straightforward way.

4.3 A realization OT approach to blocking and extended exponence

In this section I show that our realization OT model provides a unified approach to both blocking of inflectional affixes and extended morphological exponence without recourse to either a distinction between primary and secondary exponents (Noyer 1992, 1997) or multiple rule blocks (Stump 2001).

4.3.1 Tamazight Berber

I revisit the Tamazight Berber verbal morphology which is discussed in detail in both Noyer (1992, 1997) and Stump 2001, which criticize each other. I show that the single device *FEATURE SPLIT captures both blocking and extended exponence in this Afroasiatic dialect.

Noyer (1992, 1997) takes a rule-based realization approach to the Tamazight Berber verbal morphology whose paradigm is shown as follows:

(11) Completive paradigm of Tamazight Berber DAWA ‘cure’ (Abdel-Massih 1971: 171, Noyer 1992: 132, Stump 2001: 157)

		Singular	Plural
1		<i>dawa-γ</i>	<i>n-dawa</i>
2	masc	<i>t-dawa-d</i>	<i>t-dawa-m</i>
	fem	<i>t-dawa-d</i>	<i>t-dawa-n-t</i>
3	masc	<i>i-dawa</i>	<i>dawa-n</i>
	fem	<i>t-dawa</i>	<i>dawa-n-t</i>

Noyer’s analysis of the Tamazight Berber paradigm in (11), which is cited in Stump (2001: 157), is given in (12):

(12)	Rule of affixation	is a primary exponent of	is a secondary exponent of	bleeds
a.	<i>n-</i>	{1, pl}		(b), (h)
b.	<i>-γ</i>	{1}		
c.	<i>t-</i>	{2}		
d.	<i>-m</i>	{pl, masc}	{2}	(h)
e.	<i>i-</i>	{sg, masc}		
f.	<i>t-</i>	{sg, fem}		(i)
g.	<i>-d</i>	{sg}	{2}	
h.	<i>-n</i>	{pl}		
i.	<i>-t</i>	{fem}		

Noyer presents a realization theory based on *feature discharge*. Under this theory, once a morphosyntactic feature value is discharged or spelled out by an exponent, it will no longer be available for further realization. Thus, the prefix *n-* {1, pl} blocks the suffix *-γ* {1} because the first person feature value that is realized by *n-* is no longer available for realization by *-γ*. This is a case of what Noyer calls *discontinuous bleeding* in which both the blocking and blocked affixes belong to distinct position classes. Similar analyses apply to cases in which *n-* {1, pl} blocks *-n* {pl}, *-m* {pl, masc} blocks *-n* {pl}, and *t-* {sg, fem} blocks *-t* {fem}. Noyer's theory is completely reliant on Pāṇini's Principle, which requires an affix with more specific morphosyntactic or semantic content to preempt others with less specific content.

Tamazight Berber, however, has cases of extended exponence in which a morphosyntactic or semantic feature value is realized by more than one form. On Noyer's analysis, the second person plural masculine exponent *-m* co-occurs with the second person exponent *t-* (*t-dawa-m*). Additionally, the second person singular exponent *-d* co-occurs with *t-* {2} (*t-dawa-d*). These cases of extended exponence pose a challenge for a theory based on feature discharge, because if the second person feature value is first realized by *-m* {2, pl, masc} or *-d* {2, sg}, it should no longer be available for realization by *t-* {2} and therefore we should *not* expect the co-occurrence of *t-* with *-m/-d*.

In order to allow extended exponence, Noyer introduces extra machinery, i.e. distinction between primary and secondary exponents. An affix which realizes a morphosyntactic feature value as a secondary exponent presumes another affix which realizes the same morphosyntactic feature value as a primary exponent. Only an affix that realizes a morphosyntactic feature value as a primary exponent can block or be blocked by another affix which also expresses the same feature value as a primary exponent. An affix that realizes a morphosyntactic feature value as a secondary exponent *cannot* block or be blocked by another affix that expresses the same feature value as either a primary or secondary exponent. Extended exponence in Noyer's framework therefore involves occurrences of both a primary and secondary exponent. Noyer assumes that in Tamazight Berber *t-* is a primary exponent of the second person feature value which can be further realized by *-m* or *-d* which is a secondary exponent of {2} even if {2} has been discharged by *t-*.

Stump 2001 criticizes Noyer's (1992) treatment of extended exponence by pointing out several problems. He argues that "Noyer's analysis of Berber subject agreement does not embody a viable approach to inflectional morphology" (Stump 2001: 160). Stump first points out that Noyer's analysis of the Tamazight Berber verbal paradigm should be modified. He remarks that "[i]t is ... critical to Noyer's analysis that the property sets {PER[SON]: 2, NUM[BER]: sg, GEN[DER]: fem} and {PER: 2, NUM: sg, GEN: masc} be excluded" Stump (2001: 160). Notice that the feature value set {2, sg, masc} would otherwise be realized by both *t-* {2} and *i-* {sg, masc}. The prefix *i-* {sg, masc} which discharges singular would then block the occurrence of *-d* {sg (2)} which realizes singular as a primary exponent (and 2nd person as a secondary exponent). But *t-dawa-d* is a grammatical form of {2, masc, sg} while **t-i-dawa* is ungrammatical. Similarly, the feature value set {2, sg, fem} would otherwise be realized by both *t-* {2} and *t-* {sg, fem}. The prefix *t-* {sg, fem} would then block the occurrence of both *-d* {sg (2)} and *-t* {fem}. But *t-dawa-d* is a grammatical form of {2, fem, sg} while **t-t-dawa* is ungrammatical. "Noyer assumes that this is not simply an accident, but is a reflection of the fact that there simply is no gender distinction in the second-person singular of Berber verbs — that the property sets {PER: 2, NUM: sg, GEN: fem} and {PER: 2, NUM: sg, GEN: masc} are ill formed in Berber" (Stump 2001: 159). Stump argues against Noyer's assumption by pointing out that:

[Noyer's] assumption that gender is not a distinctive property in second-person singular in Berber verbs is questionable. Typologically, a system which distinguished gender in the second-person plural but not in the second-person singular would be quite unusual. And in Berber, in fact, gender is formally distinguished in 2sg pronominal-object suffixes for verbs and prepositions, in possessive suffixes for nouns, and in the system of free pronouns (Bentolila 1981: 74f); it is only with respect to subject agreement that the gender distinction fails to receive formal expression. This suggests that the identity of the 2sg forms in [(11)] is simply an accident of the rule system — a consequence of the fact that 2sg subject agreement is expressed by rules which happen not to be sensitive to gender. (Stump 2001: 160)

Additionally, Stump 2001 points out that the nasal in *t-dawa-n-t* is underlyingly /m/ which "assimilates to the place of articulation of the following *-t*" (p.161). He says that "[i]f a masculine nominal ends in *m*, the circumfixation of *t-...-t* invariably induces the assimilation of *m* as *n*; thus, *asMam* 'bitter (masc)' gives rise to *t-asMan-t* 'bitter (fem)' (Bentolila 1981: 25)" (Stump 2001: 161). Stump thus concludes that *-m* should be analyzed as {2, pl} rather than {2, masc, pl}.

Stump (2001: 161) admits that "neither of these empirical points is an argument against Noyer's theoretical assumptions: it's perfectly possible to modify his rule system in such a way as to take account of these empirical points without abandoning

his theoretical ground plan. One imaginable reworking of his analysis would be as [follows].”

(13)	Rule of affixation	is a primary exponent of	is a secondary exponent of	bleeds
a.	<i>n-</i>	{1, pl}		(b)
b.	<i>-γ</i>	{1}		
c.	<i>t-</i>	{2}		
d.	<i>-n</i>	{3, pl}		
e.	<i>i-</i>	{3, sg, masc}		(f)
f.	<i>t-</i>	{3, sg}		
g.	<i>-m</i>	{pl}	{2}	
h.	<i>-d</i>	{sg}	{2}	
i.	<i>-t</i>	{fem}	{pl}	

Stump 2001, however, argues that “there are good theoretical grounds for rejecting any such analysis [based on a distinction between primary and secondary exponents]” (p.162). Stump presents his crucial counterargument against a distinction between primary and secondary exponents based on the data from Swahili. He argues that the prefix *ha-* is a primary exponent of negation in future forms (e.g. *tu-ta-taka* ‘we will want’, *ha-tu-ta-taka* ‘we won’t want’). Since the past tense negation marker *ku-* (in contrast to the past tense positive marker *li-*, e.g. *tu-li-taka* ‘we wanted’) is closer to a verbal stem than the negation marker *ha-* (*ha-tu-ku-taka* ‘we didn’t want’), in Noyer’s 1992 system *ku-* should be added before *ha-*, which then needs to be a secondary exponent of negation because otherwise *ku-* would block *ha-*. Thus, *ha-* is paradoxically both a primary and secondary exponent of negation.

Additionally, Stump remarks that the notion of secondary exponence presents great difficulties for language learners because it is hard for a language learner to “determine the properties of which a rule is a primary exponent and those of which it is a secondary exponent” (Stump 2001: 165). He further shows that it is possible to treat the Tamazight Berber suffixes *-m* and *-d* in (14-15) as primary exponents instead of secondary exponents. See (14) (from Stump 2001: 165).

(14)	Rule of affixation	is a primary exponent of	is a secondary exponent of	bleeds
a.	<i>n-</i>	{1, pl}		(b)
b.	<i>-γ</i>	{1}		
c.	<i>-m</i>	{2, pl}		(d)
d.	<i>-d</i>	{2}		
e.	<i>-n</i>	{3, pl}		
f.	<i>i-</i>	{3, sg, masc}		(g)
g.	<i>t-</i>	{3, sg}		
h.	<i>-t</i>	{fem}	{pl}	
i.	<i>t-</i>	∅	{2}	

Stump 2001 presents an approach to the Tamazight Berber verbal morphology within the framework of Paradigm Function Morphology. In this model, realization rules which apply to the same affixal slot are placed in the same rule block. Within a rule block, Pāṇini's Principle is the only mechanism to determine which rule should apply (i.e. the Pāṇinian Determinism Hypothesis). Blocking is assumed to occur only within the same rule block, which corresponds to a single affixal slot, so cases of discontinuous bleeding in which prefixes block suffixes or vice versa are excluded by this framework. The key device to account for extended exponence is to refer to multiple rule blocks, i.e. co-occurring exponents which would otherwise violate Pāṇini's Principle are placed in distinct rule blocks. Stump proposes his analysis of blocking and extended exponence in Tamazight Berber as follows. (I simplify his notation.) The output from a rule block becomes an input to the following one.

(15)	Block I	[PER: 2] (X)	=	<i>tX</i>
		[PER: 3], [NUM: sg], [GEN: masc] (X)	=	<i>iX</i>
		[PER: 3], [NUM: sg] (X)	=	<i>tX</i>
		[PER: 1], [NUM: pl] (X)	=	<i>nX</i>
	Block II	[PER: 1], [NUM: sg] (X)	=	<i>Xy</i>
		[PER: 2], [NUM: pl] (X)	=	<i>Xm</i>
		[PER: 2] (X)	=	<i>Xd</i>
		[PER: 3], [NUM: pl] (X)	=	<i>Xn</i>
	Block III	[NUM: pl], [GEN: fem] (X)	=	<i>Xt</i>

Noyer (1992, 1997), on the other hand, criticizes such an analysis. Noyer (1997: 94) remarks that “[t]he discontinuous bleeding analysis requires only one block of rules whereas [a Word-and-Paradigm analysis] requires three blocks. From the point of view of learning the forms of the system, one must assume on [a Word-and-Paradigm analysis] that one must learn both the rule *and* the block it occurs in ... In contrast, the analysis I have given in [(12)] requires only that each affix be learned associated with its feature content.”

To briefly summarize, Stump and Noyer criticize each other's work. Stump 2001 argues that “Noyer's notion of feature discharge is not a satisfactory alternative to the postulation of rule blocks, since it depends on an empirically unmotivated and ultimately paradoxical distinction between primary and secondary exponents” (p.168). Noyer (1992, 1997), on the other hand, remarks that multiple rule blocks complicate learning processes.

I argue for a realization OT model and show that it provides a unified approach to both blocking and extended exponence without recourse to either a distinction between primary and secondary exponents or multiple rule blocks. The key device is the markedness constraint *FEATURE SPLIT which bans the realization of a morphosyntactic or semantic feature value by more than one form. I show that by varying rankings of *FEATURE SPLIT and constraints realizing the same

morphosyntactic or semantic feature value(s), we can readily capture both blocking and extended exponence. If *FEATURE SPLIT is outranked by constraints realizing the same morphosyntactic or semantic feature value(s), we will expect extended exponence. Otherwise, we will observe blocking of inflectional affixes. Let us reconsider the Tamazight Berber verbal paradigm in (11) (repeated in (16)).

(16) Completive paradigm of Tamazight Berber DAWA ‘cure’ (Abdel-Massih 1971: 171, Noyer 1992: 132, Stump 2001: 157)

		Singular	Plural
1		<i>dawa-γ</i>	<i>n-dawa</i>
2	masc	<i>t-dawa-d</i>	<i>t-dawa-m</i>
	fem	<i>t-dawa-d</i>	<i>t-dawa-n-t</i>
3	masc	<i>i-dawa</i>	<i>dawa-n</i>
	fem	<i>t-dawa</i>	<i>dawa-n-t</i>

We can directly translate the affixes in (12) (on Noyer’s (1992, 1997) analysis) into realization constraints *without* distinguishing primary from secondary exponents. Our realization OT grammar is presented in (17). Notice that the constraints {2, pl, masc}: *-m*, {2, sg}: *-d*, and {2}: *t-* need to outrank *FEATURE SPLIT because both *-m* {2, pl, masc} and *-d* {2, sg} can co-occur with *t-* {2} (*t-dawa-m*, *t-dawa-d*) so that the second person feature value is realized by two exponents. The ranking of *t-* {2} and *-m* {2, pl, masc} / {2, sg}: *-d* is indeterminate in that we lack evidence to show that the former is outranked by the latter, but I assume that it still conforms to the specificity condition which requires a constraint with more specific morphosyntactic or semantic content to outrank a less specific realization constraint. For a clearer presentation, I rank *FEATURE SPLIT higher than the other realization constraints simply to show that extended exponence is introduced by the constraints on the left side of *FEATURE SPLIT in (17). But actually if, for example, {1, pl}: *n-* outranks *FEATURE SPLIT, our results remain intact.

(17) {2, pl, masc}: *-m*, {2, sg}: *-d* >> {2}: *t-* >> *FEATURE SPLIT >> {1, pl}: *n-*,
 {sg, fem}: *t-*, {sg, masc}: *i-* >> {1}: *-γ*, {pl}: *-n*, {fem}: *-t*

Let us see how our realization OT grammar captures every paradigmatic cell in (16). Consider the tableau in (18a). The illicit form **n-dawa* {1, sg} is ruled out by the constraint {1}: *-γ* because *-γ* is not in the output.

(18) a. *dawa-γ* {1, sg}

<i>dawa</i> , 1, sg	2, pl, m: -m	2, sg: -d	2: t-	*FS	1, pl: n-	sg, f: t-	sg, m: i-	1: -γ	pl: -n	f: -t
☞ a. $\begin{array}{c} 1 \text{ sg} \\ \\ \textit{dawa} \\ \\ \textit{-}\gamma \end{array}$										
b. $\begin{array}{c} 1 \text{ sg} \\ \\ \textit{n-dawa} \end{array}$								*!		

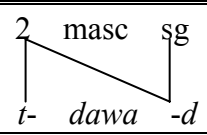
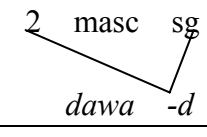
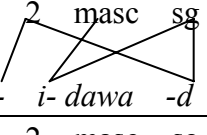
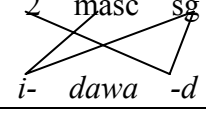
In (18b), the illicit form **dawa-γ* {1, pl} is ruled out by the constraint {1, pl}: *n-*.

b. *n-dawa* {1, pl}

<i>dawa</i> , 1, pl	2, pl, m: -m	2, sg: -d	2: t-	*FS	1, pl: n-	sg, f: t-	sg, m: i-	1: -γ	pl: -n	f: -t
☞ a. $\begin{array}{c} 1 \text{ pl} \\ \diagdown \quad \diagup \\ \textit{n-dawa} \end{array}$								*		
b. $\begin{array}{c} 1 \text{ pl} \\ \\ \textit{dawa} \\ \\ \textit{-}\gamma \end{array}$					*!					

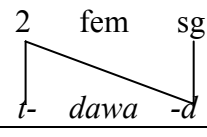
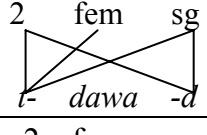
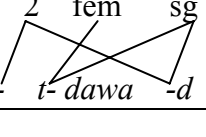
Consider the tableau in (18c). The grammatical output candidate *t-dawa-d* {2, masc, sg} wins although it violates *FEATURE SPLIT because {2} is realized by both *t-* and *-d*. The illicit form **dawa-d* {2, masc, sg} is ruled out by the constraint {2}: *t-* which outranks *FEATURE SPLIT. The illicit form **t-i-dawa-d* {2, masc, sg} is ruled out because it violates *FEATURE SPLIT twice in that not only {2} is realized by both *t-* and *-d*, but also {sg} is realized by both *i-* and *-d*. The illicit form **i-dawa-d* is ruled out by the constraint {2}: *t-*. Notice that the occurrence of the prefix *i-* {sg, masc} is ruled out in the environment of {2, masc, sg} because of both the constraints *FEATURE SPLIT which rules out **t-i-dawa-d* and {2}: *t-* which rules out **i-dawa-d*. There is absolutely no necessity that we rule out the feature value set {2, masc, sg} in the Tamazight Berber verbal morphology. Just as Stump 2001 points out, the constraint (or rule) {2}: *t-* is insensitive to gender distinction in the environment of second person subject agreement. By contrast, Noyer (1992) has to rule out the input feature value set {2, masc, sg} in (12) in order for his system to work.

c. *t-dawa-d* {2, masc, sg}

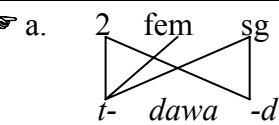
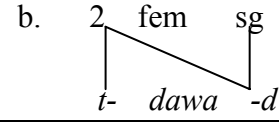
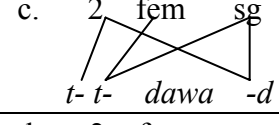
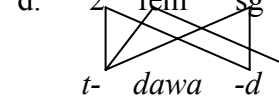
<i>dawa</i> , 2, masc, sg	2, pl, m: -m	2, sg: -d	2: t-	*FS	1, pl: n-	sg, f: t-	sg, m: i-	l: -γ	pl: -n	f: -t
a. 				*			*			
b. 			*!				*			
c. 				**!						
d. 			*!	*						

Consider the tableau in (18d). Candidate (b) is ruled out by this grammar because it violates *FEATURE SPLIT twice. Candidate (c) also violates *FEATURE SPLIT twice and is therefore ruled out. Notice that this grammar requires the prefix *t-* not to realize {2, fem, sg} since otherwise both {2} and {sg} would be realized by both *t-* and *-d*. Thus, Candidate (b) is ruled out by *FEATURE SPLIT which outranks {sg, f}: *t-*. But {sg, f}: *t-* may outrank *FEATURE SPLIT so that Candidate (b) wins. Candidate (c) which has an extra *t-* compared to Candidates (a-b) can be ruled out by a constraint banning the occurrence of phonological information that is not present in the input, i.e. DEP. See the tableau in (18e) in which {sg, f}: *t-* outranks *FEATURE SPLIT.

d. *t-dawa-d* {2, fem, sg}

<i>dawa</i> , 2, fem, sg	2, pl, m: -m	2, sg: -d	2: t-	*FS	1, pl: n-	sg, f: t-	sg, m: i-	l: -γ	pl: -n	f: -t
a. 				*		*				*
b. 				**!						*
c. 				**!						*

e. *t-dawa-d* {2, fem, sg}

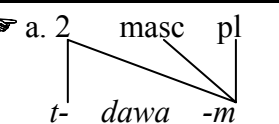
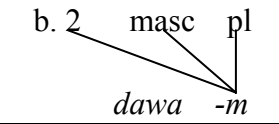
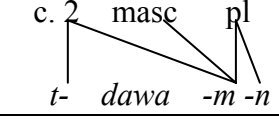
<i>dawa</i> , 2, fem, sg	2, sg: - <i>d</i>	2: <i>t</i> -	sg, f: <i>t</i> -	*FS	f: - <i>t</i>	DEP
a. 				**	*	
b. 			*!	*	*	
c. 				**	*	<i>t</i> -!
d. 				***!		- <i>t</i>

Since the form *t-dawa-d* {2, fem, sg} in which *t*- realizes {2} is homophonous to the form *t-dawa-d* {2, fem, sg} in which *t*- realizes {2, fem, sg}, the rankings of the constraints *FEATURE SPLIT and {sg, f}: *t*- will give us two equally possible options.

Again, as we can see, there is absolutely no necessity that we exclude the feature value set {2, fem, sg}. By contrast, Noyer (1992, 1997) needs to exclude this feature value set in order for his analysis in (12) to work.

Consider the tableau in (18f). The illicit form **dawa-m* {2, masc, pl} is ruled out by the constraint {2}: *t*- and the illicit form **t-dawa-m-n* {2, masc, pl} violates *FEATURE SPLIT twice because both {2} and {pl} are doubly marked.

f. *t-dawa-m* {2, masc, pl}

<i>dawa</i> , 2, masc, pl	2, pl, m: - <i>m</i>	2, sg: - <i>d</i>	2: <i>t</i> -	*FS	1, pl: <i>n</i> -	sg, f: <i>t</i> -	sg, m: <i>i</i> -	1: - γ	pl: - <i>n</i>	f: - <i>t</i>
a. 				*					*	
b. 			*!						*	
c. 				***!						

The grammatical form *t-dawa-n-t* {2, fem, pl} in which *t*- realizes {2}, *-n* realizes {pl}, and *-t* realizes {fem} is an optimal candidate chosen by our grammar.⁴ Other candidates which realize {2, fem, pl} and miss any of the above affixes will be ruled out by realization constraints. Similarly, the grammatical forms *i-dawa* {3, masc, sg} in which *i*- realizes {masc, sg}, *dawa-n* {3, masc, pl} in which *-n* realizes {pl}, and

dawa-n-t {3, fem, pl} in which *-n* realizes {pl} and *-t* realizes {fem} are all optimal candidates and do not violate *FEATURE SPLIT or any realization constraint.

Consider the tableau in (18g). The illicit form **t-dawa-t* {3, fem, sg} is ruled out by *FEATURE SPLIT because {fem} is realized by both *t-* and *-t*.

g. *t-dawa* {3, fem, sg}

<i>dawa</i> , 3, fem, sg	2, pl, m: <i>-m</i>	2, sg: <i>-d</i>	2: <i>t-</i>	*FS	1, pl: <i>n-</i>	sg, f: <i>t-</i>	sg, m: <i>i-</i>	1: <i>-γ</i>	pl: <i>-n</i>	f: <i>-t</i>
a. 3 fem sg <i>t-</i> <i>dawa</i>										*
b. 3 fem sg <i>t-</i> <i>dawa</i> <i>-t</i>				*!						

So far I have shown that we can simply translate the affixes in (12) (on Noyer's (1992, 1997) analysis) into realization constraints and the proposed OT grammar readily captures every cell of the Tamazight Berber paradigm in (11) without resorting to a distinction between primary and secondary exponents or excluding the input feature value sets {2, masc, sg} and {2, fem, sg}.

We can also use the affixes in (13), which Stump 2001 proposes to modify Noyer's original analysis of the Tamazight Berber verbal morphology. Stump's modification makes it unnecessary for Noyer's system to exclude the feature value sets {2, masc, sg} and {2, fem, sg}. On this analysis, *-m* is an underlying exponent of {2, pl} (*t-dawa-n-t*) instead of *-n*. We can directly translate the affixes in (13) into realization constraints and our proposed grammar is shown in (19). I put aside the issue of *-m* assimilating to the place of articulation of the following *-t*, which can be implemented in various ways. See Chapter 2 for a relevant discussion of the alternation of the English regular plural marker *-s* in various contexts.

(19) {2, pl}: *-m*, {2, sg}: *-d*, {fem, pl}: *-t*, {3, pl}: *-n* >> {2}: *t-* >> *FEATURE SPLIT
>> {1, pl}: *n-*, {3, sg, masc}: *i-* >> {1}: *-γ*, {3, sg}: *t-*

The grammar in (19) also captures every cell of the paradigm in (11) (repeated in (20)).

(20) Completive paradigm of Tamazight Berber DAWA ‘cure’ (Abdel-Massih 1971: 171, Noyer 1992: 132, Stump 2001: 157)

		Singular	Plural
1		<i>dawa-γ</i>	<i>n-dawa</i>
2	masc	<i>t-dawa-d</i>	<i>t-dawa-m</i>
	fem	<i>t-dawa-d</i>	<i>t-dawa-n-t</i>
3	masc	<i>i-dawa</i>	<i>dawa-n</i>
	fem	<i>t-dawa</i>	<i>dawa-n-t</i>

We can also refer to Stump’s 2001 own analysis of the Tamazight Berber verbal morphology and directly translate his realization rules in (15) into realization constraints with one small change. Notice that on Stump’s own analysis, the suffix *-d* is assumed to realize {2} instead of {2, sg}. By contrast, he analyses the suffix *-γ* as an exponent of {1, sg} instead of {1}. I analyze the suffix *-d* as an exponent of {2, sg} not only because we can avoid such analytical inconsistency in Stump 2001 by doing that, but also because otherwise it would be difficult to explain why *t-* {2} and *-d* {2} can co-occur in the environment of {2, sg} but not {2, pl}. The prefix *t-* has a wider distribution than that of the suffix *-d*, which can be captured by assigning *-d* a more specific feature value set. Our proposed grammar is shown in (21). (I put aside the issue of *-m* assimilating to the place of articulation of the following *-t*.) The grammar in (21) not only captures every cell of the paradigm in (11) (= (20)), but also avoids both analytical cyclicity and multiple rule blocks. The only difference between the grammars in (19) and (21) is that in (21) the suffix *-γ* realizes {1, sg} instead of {1} according to Stump 2001, so there is no subset relation between the constraints {1, pl}: *n-* and {1, sg}: *-γ* in (21).

(21) {2, pl}: *-m*, {2, sg}: *-d*, {fem, pl}: *-t*, {3, pl}: *-n* >> {2}: *t-* >> *FEATURE SPLIT >> {1, pl}: *n-*, {1, sg}: *-γ*, {3, sg, masc}: *i-* >> {3, sg}: *t-*

We leave our readers an exercise of confirming that both the grammars in (19) and (21) can capture every cell of the paradigm in (11) (= (20)) by drawing tableaux.

There are two issues concerning our realization OT grammars. First, Noyer (1992) proposes the Feature Hierarchy Hypothesis (22) and hypothesizes that rule ordering follows this hypothesis, i.e. a rule or affix realizing a feature value on the left precedes another rule or affix realizing a feature value on the right.

(22) 1st person > 2nd person > plural > feminine

This hypothesis is necessarily sacrificed in our OT grammar which is basically a non-cyclic model. Thus, it remains to be tested whether this hypothesis is falsifiable or indispensable in analyzing other language data. Second, some constraint rankings which do not need to conform to the specificity condition are indeterminate. For example, the ranking of {2, sg}: *-d* and {fem, pl}: *-t* is indeterminate. The issue of

indeterminacy of rule ordering can also be found in Noyer 1992, Stump 2001. For example, in (12) (on Noyer's analysis) the order of the affixes *i-* {sg, masc} and *t-* {sg, fem} is indeterminate as pointed out by Stump 2001. Stump 2001 proposes three rule blocks (15) to analyze the Tamazight Berber verbal morphology. The third rule block has only one realization rule. Each of the first two rule blocks contains four realization rules whose order is also indeterminate.

To summarize, our realization OT approach to the Tamazight Berber verbal morphology avoids both multiple rule blocks and distinction between primary and secondary exponents, so it simplifies learning processes. Assuming that *FEATURE SPLIT is inherent and part of Universal Grammar (cf. Noyer also relies on a mechanism which requires a morphosyntactic feature value not to be realized more than once; Stump relies on Pāṇini's Principle), a Tamazight Berber learner only needs to learn realization constraints (similar to realization rules) and their rankings without distinguishing primary from secondary exponents or learning how to place realization rules in several rule blocks. Additionally, our model provides a unified approach to both blocking and extended exponence by varying rankings of *FEATURE SPLIT and realization constraints. The violable device *FEATURE SPLIT underlies both blocking and extended exponence. By contrast, neither Noyer (1992, 1997) nor Stump (2001) presents a unified approach to blocking and extended exponence. To account for blocking, Noyer refers to a mechanism in the spirit of *FEATURE SPLIT, while to allow for extended exponence, he refers to a completely distinct mechanism, i.e. distinction between primary and secondary exponents. Stump 2001, on the other hand, relies on Pāṇini's Principle which only operates in one rule block to account for blocking. To account for extended exponence, he also refers to a completely distinct mechanism, i.e. multiple rule blocks. Both the rule-based realization frameworks lack a single device like *FEATURE SPLIT which can unify language phenomena such as blocking and extended exponence without distinct machinery.

4.3.2 Classical Arabic

I show that our realization OT approach also applies to the Arabic language which shows both blocking and extended exponence. Our focus is on the Classical Arabic prefixal conjugation which has been discussed in great detail in Noyer (1992, 1997). Consider the following paradigms.

(23) Classical Arabic prefixal conjugation (adapted from Noyer 1997: 4-5)

a. Imperfect

singular	dual	plural	
<i>ʔ-aktub-u</i>	<i>n-aktub-u</i>	<i>n-aktub-u</i>	1
<i>t-aktub-u</i>	<i>t-aktub-aa-ni</i>	<i>t-aktub-uu-na</i>	2, masc
<i>t-aktub-ii-na</i>	<i>t-aktub-aa-ni</i>	<i>t-aktub-na</i>	2, fem
<i>y-aktub-u</i>	<i>y-aktub-aa-ni</i>	<i>y-aktub-uu-na</i>	3, masc
<i>t-aktub-u</i>	<i>t-aktub-aa-ni</i>	<i>y-aktub-na</i>	3, fem

b. Subjunctive

singular	dual	plural	
<i>ʔ-aktub-a</i>	<i>n-aktub-a</i>	<i>n-aktub-a</i>	1
<i>t-aktub-a</i>	<i>t-aktub-aa</i>	<i>t-aktub-uu</i>	2, masc
<i>t-aktub-ii</i>	<i>t-aktub-aa</i>	<i>t-aktub-na</i>	2, fem
<i>y-aktub-a</i>	<i>y-aktub-aa</i>	<i>y-aktub-uu</i>	3, masc
<i>t-aktub-a</i>	<i>t-aktub-aa</i>	<i>y-aktub-na</i>	3, fem

c. Jussive

singular	dual	plural	
<i>ʔ-aktub</i>	<i>n-aktub</i>	<i>n-aktub</i>	1
<i>t-aktub</i>	<i>t-aktub-aa</i>	<i>t-aktub-uu</i>	2, masc
<i>t-aktub-ii</i>	<i>t-aktub-aa</i>	<i>t-aktub-na</i>	2, fem
<i>y-aktub</i>	<i>y-aktub-aa</i>	<i>y-aktub-uu</i>	3, masc
<i>t-aktub</i>	<i>t-aktub-aa</i>	<i>y-aktub-na</i>	3, fem

There are at least two interesting issues in (23). First, the second person exponent *t-* co-occurs with the second person feminine singular marker *-ii* in the environment of {2, fem, sg} (e.g. *t-aktub-ii*), which is a case of extended exponence in which {2} is realized by both *t-* and *-ii*. Second, in the environment of {3, fem, pl} we get *y-aktub-na* instead of **t-aktub-na*.

Noyer's (1997: 54) analysis of the affixes in (23) is presented as follows.

- (24)
- | | | |
|----|--------------|-------------------------|
| a. | <i>n-</i> | {1, pl} |
| b. | <i>ʔ-</i> | {1} |
| c. | <i>t-</i> | {2} |
| d. | <i>-aani</i> | {dual} |
| e. | <i>-na</i> | {pl, fem} |
| f. | <i>-uuna</i> | {pl} |
| g. | <i>-iina</i> | {fem, (2)} |
| h. | <i>t-</i> | {fem} |
| i. | <i>-u</i> | {-perfect, +indicative} |
| j. | <i>-∅</i> | jussive |
| k. | <i>y-</i> | Elsewhere |
| l. | <i>-a</i> | Elsewhere |

To explain cases like *y-aktub-na* {3, fem, pl} in which *-na* {fem, pl} blocks *t-* {f} (**t-aktub-na*), Noyer assumes that {fem} which is first discharged or realized by *-na* is no longer available for realization by *t-*. This is a case of discontinuous bleeding in which a suffix blocks a prefix. To allow cases of extended exponence such as *t-aktub-iina*, Noyer distinguishes primary from secondary exponents, i.e. *t-* is a primary exponent of {2} while *-iina* is a secondary exponent of {2}. The suffix *-iina* therefore does not block or get blocked by *t-*.

I show that by modifying Noyer's (1992, 1997) analysis of the affixes in (24), we can use our realization OT approach to account for the paradigm in (23) and readily capture both blocking and extended exponence without recourse to a distinction between primary and secondary exponents. Notice that Noyer analyzes *n-* as an exponent of {1, pl}. But *n-* can also be an exponent of {1, dual}.⁵ It seems that Noyer analyzes *n-* as an exponent of {1, pl} so that *n-* can block *-uuna*, which is analyzed as an exponent of {pl} (*n-aktub-u* vs. **n-aktub-uuna*) and he can get another case of discontinuous bleeding which he strongly advocates in his framework. In order to account for the syncretism of {1, dual} and {1, pl} forms which share the same prefix *n-*, Noyer needs to use a feature-changing mechanism to convert the feature value set {1, dual} into {1, pl}. However, Noyer (1997: 87) admits that "such [feature-changing] rules are highly costly. If alternative analyses exist, they are presumably less costly and therefore more likely to reflect speaker's knowledge of morphology. On these grounds, I will not advocate the feature-changing analysis for the Semitic forms, since I have presented what I believe to be a less costly homophony analysis." I analyze the prefix *ʔ-* as an exponent of {1, sg} and *n-* as an exponent of {1} in (23). I analyze *-aa* as an exponent of {dual}, *-uu* as an exponent of {masc, pl}, and *-ii* as an exponent of {2, fem, sg}. I analyze *-ni* as an exponent of {dual, -perf, +ind} and *-na* as an exponent of either {fem, pl} or {2, fem, sg, -perf, +ind} or {masc, pl, -perf, +ind}. I analyze *t-* as an exponent of either {2} or {3, fem} and *y-* as an exponent of {3}. I analyze *-u* as an exponent of {-perf, +ind} and *-a* as an exponent of the subjunctive mood. My analysis of the affixes in (23) is presented as follows.

- (25)
- | | | |
|----|------------|---------------------------|
| a. | <i>ʔ-</i> | {1, sg} |
| b. | <i>n-</i> | {1} |
| c. | <i>-aa</i> | {dual} |
| d. | <i>-uu</i> | {masc, pl} |
| e. | <i>-ii</i> | {2, fem, sg} |
| f. | <i>-ni</i> | {dual, -perf, +ind} |
| g. | <i>-na</i> | {fem, pl} |
| h. | <i>-na</i> | {2, fem, sg, -perf, +ind} |
| i. | <i>-na</i> | {masc, pl, -perf, +ind} |
| j. | <i>t-</i> | {2} |
| k. | <i>t-</i> | {3, fem} |
| l. | <i>y-</i> | {3} |
| m. | <i>-u</i> | {-perf, +ind} |
| n. | <i>-a</i> | {subjunctive, -ind} |

There are several potential issues concerning the above analysis. First, it is possible to analyze *-na* as a default exponent of {-perf, +ind} since it occurs in the environment of {2, fem, sg} and {masc, pl} which do not form a natural class. But what then of the suffix *-u*? The distribution of the suffix *-u* which realizes {-perf, +ind} is also irregular. It basically occurs in the contexts of both {1} and {sg} which do not form a natural class, either. If we treat both *-na* and *-u* as default exponents of {-perf, +ind}, it will be hard to explain why, for example, the exponent of {3, fem, sg, -perf, +ind} is *t-aktub-u* instead of **t-aktub-na*.⁶

Second, it is also possible to analyze, for example, *-aani* as a unitary suffix of {dual, -perf, +ind} (cf. McCarthy 2005). As Noyer (1997: 46) observes, “[w]herever the imperfect has a disyllabic suffix (*-uuna*, *-iina*, *-aani*), the subjunctive and jussive moods have only the first syllable of this suffix.” By positing a realization constraint like {dual, -perf, +ind}: *-aani*, “we are forced to assert (in effect) that it is a mere accident that the [-indicative] affixes are in all cases the first syllables of the [+indicative] affixes. I will not pursue this option” (Noyer 1997: 47). To capture this generalization, Noyer proposes a morphologically conditioned rule of truncation which says that the second syllable of a disyllabic suffix (*-uuna*, *-iina*, *-aani*) realizing {-perf, +ind} is truncated in the context of [-indicative]. By contrast, I analyze, for example, *-aa-ni* as two separate suffixes, which not only captures Noyer’s observation but also avoids a rule of truncation, which is a relatively rare and marked process.

Noyer 1997 analyzes the prefix *y-* as an elsewhere marker which is expected to appear in any prefixal position. I analyze it as an exponent of {3} since it only occurs in the context of {3}. I analyze *t-* as an exponent of {3, fem}⁷ so that based on *FEATURE SPLIT it can block *y-* in the environment of {3, fem} except where *-na* {fem, pl} occurs. It is also possible to analyze *t-* as an exponent of {fem} just as Noyer 1997 does. Then we need to explain why, for example, in the environment of {3, fem, dual} **t-y-aktub-aani* in which *t-* realizes {fem} and *y-* realizes {3} is not grammatical. There are at least two ways to rule out **t-y-aktub-aani* {3, fem, dual}. Since Classical

Arabic does not allow complex onsets of any kind, **t-y-aktub-aani* may violate a dominant phonological constraint *COMPLEX ONSET. It may also be ruled out by a morphological templatic constraint which requires a word to have only one prefix and outranks the constraint realizing *y-*. This templatic constraint may be violable. Robert Hoberman (p.c.) points out to me that Classical Arabic allows a word to have two prefixes. For example, the future tense marker *sa-* can precede *t-* {2}. See (26).

- (26) *sa-t-aqbal-uu-na-hum-aa*
 FUT-2-receive-M.PL.IND-3PL.M-DUAL
 ‘you will receive them (two)’

The jussive mood does not have an overt exponent like *-u* (imperfect) and *-a* (subjunctive). One way to express the jussive mood is to posit a zero suffix (e.g. Noyer 1997) so that we can maintain the generalization that in the so-called Classical Arabic prefixal conjugation “every verb has at least and at most one suffix and at least and at most one prefix.” (Noyer 1997: 31) We can also assume that there is simply no constraint or rule realizing the jussive mood. By doing that, we need to assume that the morphological templatic constraint requires a word to have at most one suffix rather than exactly one suffix.

Last but not least, we do need a templatic constraint⁸ to account for the distribution of *-u*, which realizes {imperfect}, and *-a*, which realizes {subjunctive}. The empirical generalization is that *-u* and *-a* show up in the positions which no other suffix can fill in. Without a templatic constraint which requires a word to have at most one suffix, it will be hard to explain why *t-aktub-na* {2, fem, pl, -perf, +ind} is grammatical while **t-aktub-u-na* in which *t-* realizes {2}, *-u* realizes {-perf, +ind}, and *-na* realizes {fem, pl} is ungrammatical. This templatic constraint can be defined as follows.

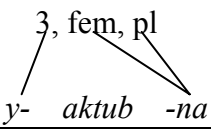
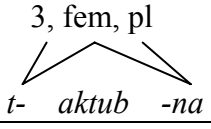
- (27) Verb Stem - Suffix (≤ 1): An inflected verb can have at most one suffix.

I directly translate the Classical Arabic affixes (25) (on my analysis) into realization constraints. The grammar which consists of realization constraints, *FEATURE SPLIT, and Verb Stem - Suffix (≤ 1) is presented in (28).⁹ For a clearer presentation, I rank *FEATURE SPLIT lower than the realization constraints which can introduce extended exponence. The grammar in (28) captures every cell of the paradigms in (23).¹⁰ The ranking in (28) is by no means the only possible ranking given that the ranking of constraints like {1, sg}: *ʔ-* and {3}: *y-* is indeterminate.

- (28) {2, fem, sg, -perf, +ind}: *-na*, {masc, pl, -perf, +ind}: *-na*, {dual, -perf, +ind}:
-ni >> {2, fem, sg}: *-ii*, {masc, pl}: *-uu*, {dual}: *-aa* >> {2}: *t-* >>
 *FEATURE SPLIT, Verb Stem - Suffix (≤ 1) >> {fem, pl}: *-na* >> {3, fem}: *t-*,
 {1, sg}: *ʔ-* >> {1}: *n-*, {3}: *y-*, {-perf, +ind}: *-u*, {subjunctive, -ind}: *-a*

Cases of extended exponence like *t-aktub-ii* {2, fem, sg, subjunctive, -ind} in which {2} is realized by both *t-* and *-ii* can be accounted for by ranking the realization constraints {2, fem, sg}: *-ii* and {2}: *t-* higher than *FEATURE SPLIT. The case of discontinuous bleeding in which *-na* {fem, pl} blocks *t-* {3, fem} in the context of {3, fem, pl} (*y-aktub-na* vs. **t-aktub-na*) can only be ascribed to *FEATURE SPLIT and the ranking *FEATURE SPLIT >> {fem, pl}: *-na* >> {3, fem}: *t-* >> {3}: *y-*. See the following tableau.

(29) *y-aktub-na* {3, fem, pl}

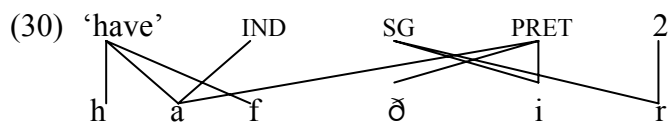
<i>aktub</i> , 3, fem, pl	*FEATURE SPLIT	{fem, pl}: <i>-na</i>	{3, fem}: <i>t-</i>	{3}: <i>y-</i>
a. 			*	
b. 	*!			*

Our realization OT model again provides a unified approach to both blocking and extended exponence in the Arabic language without recourse to a distinction between primary and secondary exponents.

We can imagine that Stump 2001 would assume that, for example, blocking of *n-* {1} by *ʔ-* {1, sg} in Arabic takes place within a single rule block while the occurrence of both *t-* {2} and *-ii* {2, fem, sg} is accounted for by placing *t-* and *-ii* in two distinct rule blocks. There are two problems with this analysis. First, blocking and extended exponence are explained by two distinct mechanisms, i.e. multiple rule blocks (for extended exponence) and Pāṇini's Principle (for blocking), which is assumed to operate within one rule block. By contrast, the violable markedness constraint *FEATURE SPLIT unifies both blocking and extended exponence. Second, as Noyer (1992, 1997) argues, cases of discontinuous bleeding pose a challenge for a framework like Stump 2001 in which blocking is assumed not to take place among rule blocks or position classes. By contrast, the markedness constraint *FEATURE SPLIT is especially good at accounting for cases of discontinuous bleeding in which a blocking and blocked affix belong to different positions.

4.3.3 More examples of extended exponence

Apart from Tamazight Berber and Arabic, cases of extended exponence occur cross-linguistically. Just to give a few examples. In Icelandic, for example, the verb *hafðir* {2, sg, preterite, indicative} “can be regarded as having five constituent elements that contribute to its phonological shape: the stem /h(a)f/, the vowel of this stem (a), and the suffixed elements /-ð/, /-i/, and /-r/” (Anderson 1992: 55). See (30).



As we can see, the preterite is realized by the verbal stem, the suffixes *-ð* and *-i*. The singular is realized by both *-i* and *-r*.

Consider the Welsh Romany verbal paradigm (Stump 1993: 450) in (31). As we can see, the preterite is realized by both *-d* and the suffixes which also realize person and number.

(31)	Present	Imperfect	Preterite
1sg	<i>kamáva</i>	<i>kamávas</i>	<i>kamdóm</i>
2sg	<i>kamésa</i>	<i>kamésas</i>	<i>kamdán</i>
3sg	<i>kaméla</i>	<i>kamélas</i>	<i>kamdás</i>
1pl	<i>kamása</i>	<i>kamásas</i>	<i>kamdám</i>
2pl	<i>kaména</i>	<i>kaménas</i>	<i>kamdán</i>
3pl	<i>kaména</i>	<i>kaménas</i>	<i>kamdé</i>

Matthews (1991: 180) points out that “in Ancient Greek, Perfective has extended exponents in *e-le-lý-k-e-te* (*le-*, *y* not *y:*, *-k-*); likewise Past (*e-*, *-e-*); likewise Active (*-k-*, *-e-*, *-te*). In this whole word only Indicative has a simple exponent (*-e-*), and in that position there is still overlapping.”

Harris 2007 shows that in Tsova-Tush, a language of the Nakh-Dagestanian language family, markers of gender-number agreement which are traditionally called class markers (CM) can appear several times (up to five times) in a single verb. See (32) (from Harris 2007). The agreement in (32) is conditioned by the absolutive form *c'a*.

(32)	<i>tišⁿ</i>	<i>c'a</i>	<i>daḥ</i>	<i>d-ex-d-o-d-anõ</i>
	old	house (d/d).ABS	PREVERB	CM-destroy-CM.TRANS-PRESENT-CM-EVIDI
	‘They are evidently tearing down the old house.’			

Similar cases can be found in Archi, another language of the Nakh-Dagestanian language family. Consider the example in (33) (Kibrik 1977 cited in Corbett 1991: 108 and Harris 2007). Class markers representing the class of the head, which is not in the example, show up several times in a word. The root is *as:á* ‘of myself’, while *t:u* is a suffix that form adjectives.

(33)	<i>d-as:á-r-ej-r-u-t:u-r</i>
	II-of.myself-II-SUFFIX-II-SUFFIX-SUFFIX-II
	‘my own’ [female]

The above examples show that extended exponence occurs cross-linguistically.^{11,12} These cases of extended exponence are fully compatible with our realization OT model which relies on violable markedness constraints like *FEATURE SPLIT. It is no surprise to the proposed realization OT model that in some cases of extended exponence, a morphosyntactic or semantic feature value can be realized by several exponents.

4.4 Alternative approaches to blocking and extended exponence

In this section, I discuss several morphological models with respect to blocking and extended exponence. In section 4.3, I have shown that our realization OT model provides a unified approach to both blocking and extended exponence without recourse to either multiple rule blocks (Stump 2001) or a distinction between primary and secondary exponents (Noyer 1992, 1997). Productive cases of extended exponence can be readily captured by ranking *FEATURE SPLIT lower than constraints realizing the same morphosyntactic or semantic feature value(s).

By contrast, in an OT model in which inflectional affixes are introduced through inputs (e.g. Kurisu 2000, Bonet 2004) one can imagine that cases of blocking and extended exponence need to be stipulated in the input. For example, in Classical Arabic the second person marker *t-* can co-occur with the second person feminine singular marker *-ii* (*t-aktub-ii*). In such an OT model, the co-occurrence of *t-* and *-ii* will simply be stipulated in the input (*/t-aktub-ii/*). On the other hand, the blocking of *t-* {3, fem} by *-na* {fem, pl} in Classical Arabic also needs to be stipulated in the input and can conceivably be expressed as either */y-aktub-na/* or */y-aktub + {-na > t-}/*.

Incremental morphological models (Lieber 1992, Steele 1995, Wunderlich 1996, Wunderlich and Fabri 1996)¹³ in which affixes are assumed to introduce morphosyntactic or semantic feature values exclude the possibility of extended exponence and especially cases in which a morphosyntactic feature value is realized by several exponents whose morphosyntactic feature value sets are the same or among whose morphosyntactic feature value sets there exists a subset relation because these models assume that affixation is strictly information-adding. Stump (2001: 4) remarks that “incremental theories deny that instances of extended exponence actually arise, and must therefore resort to extraordinary means to accommodate those that do.”

4.5 Conclusion

In this chapter I show that the proposed realization OT model provides a unified approach to both blocking and extended exponence without recourse to either a distinction between primary and secondary exponents (Noyer 1992, 1997) or multiple rule blocks (Stump 2001). I focus my discussion on Tamazight Berber and Classical Arabic conjugations which have been widely discussed in the literature. I show that by ranking *FEATURE SPLIT lower than constraints realizing the same morphosyntactic or semantic feature values we observe extended exponence; otherwise, we observe blocking. The markedness constraint *FEATURE SPLIT, which bans the realization of a morphosyntactic or semantic feature value by more than one form, is especially good at accounting for cases of discontinuous bleeding (Noyer 1992, 1997) in which

blocking takes place among different positions. Compared to the proposed realization OT model, an OT model in which affixes are introduced through inputs needs to make bald stipulations to deal with both blocking and extended exponence. Incremental morphological models in which affixes introduce morphosyntactic or semantic feature values must be energetic in excluding extended exponence and especially cases in which a morphosyntactic or semantic feature value is realized by several exponents whose feature content is the same or among whose feature value sets there exists a subset relation because these incremental models assume that affixation is strictly information-adding. But if some day they admit the existence of cases of extended exponence like those in Tamazight Berber and Classical Arabic, they need to resort to extraordinary machinery to account for them.

Notes

¹ Stump's (2001) Paradigm Function Morphology differs from Anderson's (1992) A-Morphous Morphology in that in the latter framework "realization rules belonging to the same block are assumed to be linearly ordered, and competition among members of the same block is in all instances resolved by a principle of disjunctive ordering (according to which the application of the first competitor precedes and excludes that of all subsequent competitors)" (Stump 2001: 58).

² The constraint PRIORITY is a purely stipulative constraint whose nature is not clear. See Chapter 2 for a relevant discussion.

³ In addition to the auxiliary verb in (10), Haspelmath 1993 lists 17 verbs in which the negation markers *t(A)-* and *-č* are in complementary distribution and attach to main verbs. Uslar 1896 lists about 60 such verbs in an earlier stage of the Lezgian language. Most Lezgian verbs have both synthetic and periphrastic negation structures. The negation markers *t(A)-* and *-č* are in complementary distribution and never co-occur in the same negation structure. The suffix *-č* realizes [+ind] and attaches to main verbs, while the elsewhere negation marker *t(A)-* occurs in a periphrastic structure and is prefixed to the auxiliary AWUN 'do'.

⁴ The order of the suffixes *-n* (or */-m/*) and *-t* may be determined by phonology. The word-final cluster [nt] is more optimal than the final cluster [tn] because the former satisfies the Sonority Hierarchy Principle (Kenstowicz 1994) which requires a coda to have a falling sonority contour, although various types of consonants can be underlyingly adjacent in Tamazight Berber (see Abdel-Massih 1971). We can also use a constraint to make [nt] win over [t^ən] which has a "transitional vowel", though such a constraint to rule out a transitional vowel needs to be ranked low in Tamazight Berber because transitional vowels often occur to break up complex consonant clusters.

⁵ I am grateful to Robert Hoberman for pointing out this to me. Noyer 1997 simply puts an asterisk * in the slot of {1, dual}.

⁶ Robert Hoberman (p.c.) points out to me that *-na* can be analyzed as an exponent of {-perf, +ind} which occurs after long and high vowels. This analysis will also work out. I put it aside since phonological contexts are not introduced to analyze other exponents in (23).

⁷ If we adopt Stump's 2001 idea that there is no such thing as context, and everything is content, then *t-* is better analyzed as an exponent of {3, fem}. Since *t-* only shows up in the slots of {3, fem}, its content must be {3, fem}.

⁸ Noyer 1997 also uses a morphological template to account for the Arabic inflectional system.

⁹ It is crucial that the constraint Verb Stem - Suffix (≤ 1) outrank the constraints realizing *-u* and *-a*. The ranking of Verb Stem - Suffix (≤ 1) and other constraints could be adjusted. Robert Hoberman (p. c.) points out to me that the suffixes *-u* and *-a* can be followed by pronominal object markers, which will not violate the templatic constraint Verb Stem - Suffix (≤ 1) if these pronominal object markers are clitics. Below are the arguments from Bob Hoberman that the pronominal object markers are clitics, not suffixes. "(1) The same forms mark pronominal objects of verbs (all tenses), possessors of nouns, and objects of prepositions: *yaktubu=haa* 'he writes it (3f.sg)', *kataba=haa* 'he wrote it', *baytu=haa* 'her house', *min=haa* 'from her'. (2) There is next to no phonological interaction between the object markers and the base, while the subject markers interact more significantly with the verb base. The simple phonological interactions that do exist between the object markers and the verbal base are identical whether the base is a noun, a verb, or a preposition. (3) To place focus on the object pronoun, it can be detached from the verb and attached to the pseudo-preposition *ʔiyyaa=*, which has no other function: *ʔiyyaa=haa yaktubu* 'he writes it' or 'it is what he writes'" (Robert Hoberman p. c.).

¹⁰ There are several potential ways to derive the order of the suffixes *-ii*, *-uu*, *-aa*, *-na*, and *-ni*. We can simply use a templatic constraint to require, for example, *-na* to follow *-ii*, which is a stipulation. The scope constraint does not seem to work here. According to Bybee 1985, aspect markers should be closer to the verbal stem than person and number markers. Since *-ni* and *-na* are aspect markers, they should be closer to the verbal stem than *-ii*, *-uu*, and *-aa*, which are number and/or person markers. But we observe the opposite order. The order of these suffixes may arise because of phonotactic constraints. Forms like *-ii-na*, *-uu-na*, and *-aa-ni* are phonologically well-formed in contrast to **-na-ii*, **-na-uu*, and **-ni-aa* given that "[h]iatus is intolerable ... because ONSET is undominated in Arabic" (McCarthy 2005: 187). Strategies to repair these illicit forms such as consonant insertion and vowel deletion may be more costly than simply placing, for example, *-na* after *-ii* in the proposed realization OT model.

¹¹ Anderson 1986 tries all means to deny cases of extended exponence, especially cases in which a morphosyntactic feature value is realized by several exponents among whose morphosyntactic feature value sets there exists a subset relation, because these cases pose a serious problem for Pāṇini's Principle, which his framework is centered on. Anderson 2001, however, admits that "multiple formal realization of the same inflectional content does indeed occur in natural language" (p.1).

¹² Extended exponence arguably occurs in Germanic languages such as English, German. For example, the past tense form of the verb *sell* is *sold* which arguably consists of both a past tense stem *sol-* and a regular past tense suffix *-d*. Similar examples can be found in German (e.g. *Gast* (singular), *Gäste* (plural) 'guests'). See Matthews 1991 for a detailed discussion. Based on Sympathy Theory (McCarthy 1999) which can be incorporated into our realization OT model, Kurisu 2000 analyzes German plural nouns which are both suffixed and umlauted within an OT model in which affixes are introduced through inputs. By contrast, Clahsen 1999 shows that these forms are learned as wholes; as such they present no discernible problem for any theoretical model.

¹³ Wunderlich 1996 presents a lexical incremental morphological model in which affixes are assumed to introduce morphosyntactic feature values. He terms this model "Minimalist Morphology". See also Wunderlich and Fabri 1996. Wunderlich 2001, however, presents an OT approach to the Dalabon inflectional morphology under the tenet of Minimalist Morphology. Under this approach, the input consists of abstract morphosyntactic feature values while the output consists of affixes that realize these abstract morphosyntactic feature values. Wunderlich's 2001 grammar contains no realization constraint. As we can see, Wunderlich 2001 actually presents a lexical realizational model. It is thus not clear how Minimalist Morphology within the framework of Optimality Theory comprises both an incremental model in which morphosyntactic feature values are introduced by affixation and a realizational model in which morphosyntactic feature values license affixation.

Chapter Five

A realization OT approach to full and partial identity of forms

5.1 Introduction

In this chapter I present a realization OT account of full and partial identity of forms, i.e., paradigmatic syncretism and cases in which lexemes share the same inflectional formative. My approach involves both output-to-output correspondence constraints (Benua 1995, McCarthy and Prince 1995, Kenstowicz 1997, Kager 1999) and realization constraints that associate morphosyntactic or semantic feature values with phonological forms. I show that this approach has advantages over rule-based accounts such as feature impoverishment-plus-feature insertion (Noyer 1998) and rules of referral (Zwicky 1985, Stump 1993) in that the constraint-based approach provides a unified account of both full and partial identity of forms within inflectional morphology.

The organization of this chapter is set as follows. In section 5.2 I briefly discuss Noyer's 1998 approach to syncretism based on a feature-impoverishment-plus-feature-insertion theory. I show that divergent bidirectional syncretism (DBS) (Baerman 2004, Baerman, Brown, and Corbett 2005) poses a problem for Noyer's theory and DBS can be captured by both rules of referral and output-to-output correspondence constraints. In section 5.3 I show that output-to-output correspondence constraints have advantages over rules of referral in that the former can also account for partially identical inflected forms and therefore have a wider scope of application than rules of referral. I conclude in section 5.4.

5.2 Paradigmatic syncretism

Syncretism is a hot topic in inflectional morphology. It refers to a phenomenon in which several paradigmatic slots share the same form. There are two common ways to capture synchronic syncretism under realization models. First, identical paradigmatic slots can be assumed to share the same morphosyntactic feature value set, which naturally explains why these slots are phonologically identical (Stump 1993, Embick and Halle 2004, Wunderlich 2005, among many others). For example, in Hupa, an Athabascan language, the first and second person plural object markers are identical (Golla 1970, Embick and Halle 2005). Consider the paradigm in (1) (from Embick and Halle 2005).

(1) Hupa Subject / Object Markers

	Subject	Object
1sg	<i>W-</i>	<i>Wĭ-</i>
2sg	<i>n-</i>	<i>nĭ-</i>
1pl	<i>dĭ-</i>	<i>noh-</i>
2pl	<i>oh-</i>	<i>noh-</i>

As we can see, the {1, pl} and {2, pl} object marker *noh-* share the same feature value {pl}. An economical way to capture this syncretism as argued by Embick and Halle 2005 would be to refer to underspecification of feature values and assume that *noh-* only realizes {pl} so that it can be in both the {1, pl} and {2, pl} slots.

The first approach to syncretism based on underspecification of feature value sets is employed in various realizational morphological models. A second way to capture syncretism is to refer to a mechanism which connects two paradigmatic cells with distinct feature value sets. For example, Stump 1993 proposes an approach to Macedonian syncretism based on rules of referral (Zwicky 1985). Consider the Macedonian partial paradigm in (2).

(2) Macedonian partial verbal paradigm (adapted from Stump 1993: 452)

	Imperfect			Aorist				
	I	II	III	I	II	III		
<i>padn-</i> ‘fall’								
1sg	<i>padn</i>	<i>-e</i>	<i>-v</i>	<i>padn</i>	<i>-a</i>	<i>-v</i>		
2sg	<i>padn</i>	<i>-e</i>	<i>-še</i>	<i>padn</i>	<i>-a</i>			
3sg	<i>padn</i>	<i>-e</i>	<i>-še</i>	<i>padn</i>	<i>-a</i>			
1pl	<i>padn</i>	<i>-e</i>	<i>-v</i>	<i>-me</i>	<i>padn</i>	<i>-a</i>	<i>-v</i>	<i>-me</i>
2pl	<i>padn</i>	<i>-e</i>	<i>-v</i>	<i>-te</i>	<i>padn</i>	<i>-a</i>	<i>-v</i>	<i>-te</i>
3pl	<i>padn</i>	<i>-e</i>	<i>-a</i>	<i>padn</i>	<i>-a</i>	<i>-a</i>		

As we can see, the past tense (imperfect and aorist) marker *-v* occurs in the environment of non-third person (either first or second person). The question is why *-v* does not appear in the environment of {2 sg}. Stump posits a rule of referral which says that “[i]n the past tenses, the second person singular has the same form as the third person singular” (Stump 1993: 452). This rule of referral preempts the less specific *v*-suffixation rule. Stump’s treatment of the Macedonian syncretism is criticized by Bobaljik 2001 which advocates a feature-impoverishment theory (Noyer 1997, 1998) under which the second person feature value is deleted in the environment of the past tense so that the form of {2 sg} syncretizes with the default form which is the form of third person. Bobaljik remarks that rules of referral are not restrictive about syncretic directions and by contrast, “[t]he impoverishment rule ... [assumes] that third person is a default (either in terms of the rules of exponence in Macedonian or universally)” (Bobaljik 2001: 63).

Since underspecification of feature values is much less controversial, I focus my discussion on phenomena whereby either a rule of referral or feature-impoverishment

(-plus-insertion) is supposed to apply. I show that divergent bidirectional syncretism (DBS) poses a problem for a feature-impooverishment-plus-insertion theory (Noyer 1998) and DBS can be captured by both rules of referral and output-to-output correspondence constraints in our realization OT model.

5.2.1 Divergent bidirectional syncretism

Noyer 1998 makes a strong empirical claim that under the impoverishment-plus-insertion theory, systematic syncretisms “will always move from a more marked to a less marked state” (p.282). According to this theory, when two paradigmatic cells are directionally syncretic, it is always the more marked feature value set that is converted into the less marked one.

Divergent bidirectional syncretism (DBS) (Baerman 2004, Baerman, Brown, Corbett 2005) poses a problem for this empirical claim. Baerman (2004: 816) gives the following definition: Under divergent bidirectional syncretism, there is a feature value x that takes the form associated with feature value y in some contexts, while in other contexts y takes the form associated with x . Baerman illustrates DBS with cases from the Latin second declension, Classical Arabic declension, and Diyari declension.¹

Consider the Latin second declension. The suffix *-us* is the exponent of the nom(inative) singular (sg) and marks the nom sg of both default masculine nouns and a group of neuter nouns including *vulgus* ‘crowd’, *vi:rus* ‘poison’, and *pelagus* ‘sea’. By contrast, *-um* is the exponent of the acc sg and marks the acc sg of both default neuter and default masculine nouns. See (3).

(3) The Latin second declension (adapted from Baerman 2004: 816)

	DEFAULT NEUTER	DEFAULT MASCULINE	NOM & ACC in <i>-us</i>
	‘war’	‘slave’	‘crowd’
NOM SG	<i>bell-um</i>	<i>serv-us</i>	<i>vulg-us</i>
ACC SG	<i>bell-um</i>	<i>serv-um</i>	<i>vulg-us</i>
GEN SG	<i>bell-ī</i>	<i>serv-ī</i>	<i>vulg-ī</i>
DAT SG	<i>bell-ō</i>	<i>serv-ō</i>	<i>vulg-ō</i>
ABL SG	<i>bell-ō</i>	<i>serv-ō</i>	<i>vulg-ō</i>

The acc sg of nouns such as *vulgus* ‘crowd’ syncretizes with the nom sg by taking *-us* as its exponent. An analysis based on impoverishment-plus-insertion will delete the acc feature value and add the nom feature value so that the vocabulary item *-us* ↔ nom sg can be inserted, as in (4).² This analysis conforms to the tenet of impoverishment-plus-insertion that the form of a less marked feature value always prevails.

- (4) a. acc sg → sg → nom sg (in the environment of nouns like *vulgus*)
 b. *-us* ↔ nom sg

The syncretism of the nom sg and acc sg of default neuter nouns, however, contradicts the tenet of impoverishment-plus-insertion. The nom sg of the second declension default neuter nouns takes on the form of the acc sg. Given that nom is universally less marked than acc (see e.g., Comrie 1975, 1976, Woolford 2001), impoverishment-plus-insertion unexpectedly moves from a less marked to a more marked state:

- (5) a. nom sg → sg → acc sg (in the environment of default neuter nouns)
 b. *-um* ↔ acc sg

Another case of DBS comes from Classical Arabic declension. According to Baerman, “(i)n the so-called sound plurals (formed by suffixation), genitive and accusative are syncretic, marked by the ending *-i:*, which corresponds to the distinct genitive of the default type. Diptotic nouns (certain adjectival stems, some broken plurals, and some personal names)³ likewise have a syncretic genitive/accusative, but the ending is *-a*, corresponding to the distinct accusative of the default type” (p.817). As we can see from (6), the gen(itive) of diptotic nouns takes on the form of the acc. By contrast, the acc of sound plurals takes on the form of the gen.

- (6) Classical Arabic declension (adapted from Fischer 1997: 196 and Baerman 2004: 817)⁴

	PLURAL 'believers.PL'	TRIPTOTIC (DEFAULT) PATTERN 'believer' 'black one'		DIPTOTIC 'black one'
NOM	<i>muʔmin-u:</i>	<i>muʔmin-u</i>	<i>ʔaswad-u</i>	<i>ʔaswad-u</i>
GEN	<i>muʔmin-i:</i>	<i>muʔmin-i</i>	<i>ʔaswad-i</i>	<i>ʔaswad-a</i>
ACC	<i>muʔmin-i:</i>	<i>muʔmin-a</i>	<i>ʔaswad-a</i>	<i>ʔaswad-a</i>

According to Comrie (1975, 1976), acc is universally less marked than gen. (See the Case Hierarchy in (7).) Therefore, it is against the tenet of impoverishment-plus-insertion that the acc of sound plurals takes the form of the gen, a more marked feature value.

- (7) The Case Hierarchy (Comrie 1975, 1976)
- | | | | | | | |
|---------|---|---------------|---|-----------------|---|---------|
| subject | > | direct object | > | indirect object | > | oblique |
| (nom) | | (acc) | | (dative) | | (gen) |

The third instance of DBS comes from Diyari declension. In Diyari, the abs(olutive) case has a zero exponent and the suffix *-na* is the exponent of the acc. As

we can see from (8), the abs of Type V nouns (male personal names) takes on the marker of the acc, i.e., a less marked feature value takes on the form of a more marked one, given that Diyari is an ergative language. This again violates the tenet of impoverishment-plus-insertion.

(8) Diyari declension (adapted from Austin 1981: 47-50, 61, Baerman 2004: 818)

	I	II	III	IV	V
ERG	-(<i>ya</i>) <i>li</i>	- <i>li</i>	- \emptyset	- <i>ndu</i>	- <i>li</i>
ABS	- \emptyset	- \emptyset	- \emptyset	- <i>ni</i>	- <i>ḥa</i>
ACC	- \emptyset	- <i>ḥa</i>	- <i>ḥa</i>	- <i>ḥa</i>	- <i>ḥa</i>

I = singular nouns

II = non-singular nouns, non-singular 3rd person pronouns, singular pronouns

III = non-singular 1st and 2nd person pronouns

IV = female personal names, singular pronouns

V = male personal names

In effect, as long as there is a markedness difference between the two feature values *x* and *y* in a case of DBS, it will pose a potential problem for the empirical claim that syncretism obeys markedness.

One may try to save this empirical claim by assuming that the form of a marked feature value acts as a default marker. For example, in Latin, *-um* may be treated as a default marker which appears elsewhere. To account for the syncretism of the nom sg and acc sg of default neuter nouns, impoverishment-plus-insertion will delete the nom feature value so that the default marker *-um* can be inserted:

(9) a. nom $\rightarrow \emptyset$ / default neuter

b. *-um* \leftrightarrow Elsewhere

The same analysis applies to the syncretism of the gen and acc of sound plurals in Classical Arabic. That is, the gen exponent *-i* is treated as a default. The acc feature value is deleted in the environment of sound plurals so that *-i* can be inserted (ignoring the vowel lengthening of the plural marker for the moment). The syncretism of the abs and acc in Diyari can be analyzed in the same way. The acc exponent *-ḥa* is a default marker. The abs feature value is deleted in the environment of male personal names so that *-ḥa* can be inserted.

It is, however, unmotivated to assume that the form of a marked feature value acts as a default in the cases of DBS in question. Bobaljik 2001 argues in favor of the impoverishment theory in Noyer (1997, 1998) and implicitly suggests that the form of

a universally less marked feature value tends to be a default. Bobaljik criticizes Stump's (1993) account of the syncretism of the second person singular and third person singular past tense verb forms in Macedonian and remarks that rules of referral are not restrictive about syncretic directions. Bobaljik 2001 says that by contrast, "[t]he impoverishment rule ... [assumes] that third person is a default (either in terms of the rules of exponence in Macedonian or universally)" (p.63). Therefore, *-us* ↔ nom sg should be a more suitable candidate for a default than *-um* ↔ acc sg in the Latin second declension; *-a* ↔ acc should be more suitable for a default than *-i* ↔ gen in Classical Arabic; and *-∅* ↔ abs should be more suitable for a default than *-na* ↔ acc in Diyari, because the former feature values are universally less marked than the latter ones, respectively. Additionally, within these languages it is not clear why we should choose the forms of the latter feature values as defaults rather than those of the former ones, given that the forms of both marked and less marked syncretic feature values occupy equal numbers of paradigmatic cells as we can see in (3),⁵ (6), and (8).

To briefly summarize, the above cases of DBS pose a problem for the strong claim made under the impoverishment-plus-insertion theory that the form of a less marked feature value always prevails. To account for cases of DBS, impoverishment-plus-insertion needs to introduce the form of a less marked feature value in some cases and the form of a more marked feature value in others, or it sometimes needs to assume an unmotivated default marker.

5.2.2 A realization OT account of paradigmatic syncretism

In this section I present a realization OT account of the above-mentioned cases of directional syncretism and propose the constraint ranking schema in (10).

(10) output-to-output correspondence constraints >> realization constraints

Output-to-output (OO) correspondence constraints (Benua 1995, McCarthy & Prince 1995, Kenstowicz 1997, Kager 1999) make a new form identical to a base form. These constraints are "asymmetrical" because there is a direction between a base and a copier. It is always a base that determines the phonological shape of a copier and not vice versa. Since these constraints reflect a directional copying process, they can be well applied to cases of DBS in which there is a clear syncretic direction.

I propose two relevant OO correspondence constraints⁶ and two realization constraints in (11) to account for the syncretism of the nom sg and acc sg of both default neuters and neuter nouns like *vulgus* in the Latin second declension.

(11) a. IDENT (acc sg (base), nom sg / def(a)ult n(euter)) (F): Corresponding segments of the forms of both the base acc sg and the nom sg in the context of a default neuter have identical values for any phonological feature. (IDENT AN)⁷

- b. IDENT (nom sg (base), acc sg / nouns like VULG) (F): Corresponding segments of the forms of both the base nom sg and the acc sg in the context of nouns like VULG have identical values for any phonological feature.
(IDENT NA)
- c. {nom sg}: *-us*: Nominative singular is marked by the suffix *-us* in the output.
- d. {acc sg}: *-um*: Accusative singular is marked by the suffix *-um* in the output.

Let us first consider the syncretism of the nom and acc of default neuters. I assume that an input contains both a lexical stem and inherent features of a lexeme, and abstract morphosyntactic feature values, and the function *Gen* generates an infinite list of phonological forms which spell out the lexeme and the abstract morphosyntactic feature values.⁸ I assume that, for example, an input contains both the lexeme BELL whose lexical stem is *bell* and the morphosyntactic feature values nom sg; I assume that *-um* {acc sg} is a base whose phonological form is to be copied.⁹ Consider the tableau in (12). The word *bell-um* is the winning candidate although it violates the lower ranked constraint {nom sg}: *-us*. The candidate **bell-us* fatally violates the OO correspondence constraint IDENT AN because /s/ of *-us* does not correspond to /m/ of *-um* in the base with respect to a phonological feature (e.g., voicing, nasality).¹⁰

(12) BELL (Latin)

BELL (def n), nom, sg <i>bell</i> Base: acc sg: <i>-um</i>	IDENT AN	IDENT NA	{nom sg}: <i>-us</i>	{acc sg}: <i>-um</i>
☞ a. BELL (def n) nom sg ∨ <i>bell</i> <i>-um</i>			*	
b. BELL (def n) nom sg ∨ <i>bell</i> <i>-us</i>	*!			

The same grammar can account for the syncretism of the nom sg and acc sg of nouns like VULG in the Latin second declension. I assume that the input comprises the lexeme VULG and its lexical stem *vulg* and the morphosyntactic feature values acc sg. I also assume that the relevant base is *-us* {nom sg}. The crucial output candidate **vulg-um* {acc sg} fatally violates the constraint IDENT NA because /m/ of *-um* does not correspond to /s/ of *-us* in the base with respect to phonological features such as voicing, nasality. The form *vulg-us* {acc sg} is the winning candidate despite its violation of the lower-ranked constraint {acc sg}: *-um*. See (13).

(13) VULG (Latin)

VULG, acc, sg <i>vulg</i> Base: nom sg: <i>-us</i>	IDENT AN	IDENT NA	{nom sg}: <i>-us</i>	{acc sg}: <i>-um</i>
a. VULG acc sg √ <i>vulg</i> <i>-us</i>				*
b. VULG acc sg √ <i>vulg</i> <i>-um</i>		*!		

To account for the syncretism of the gen and the acc in Classical Arabic, I propose the following constraints in (14).

- (14) a. IDENT (gen (base), acc / sound plural) (vowel height): Corresponding segments of the forms of both the base gen and the acc in the context of a sound plural have identical values for vowel height. (IDENT GA (VH))
- b. IDENT (acc (base), gen / diptote) (vowel height): Corresponding segments of the forms of both the base acc and the gen in the context of a diptote have identical values for vowel height. (IDENT AG (VH))
- c. {sound pl}: long vowel: Sound plurals (formed by suffixation) are marked by long vowels. ({sound pl}: LV)
- d. {acc}: *-a*: The accusative is marked by the suffix *-a* in the output.
- e. {gen}: *-i*: The genitive is marked by the suffix *-i* in the output.

Let us first consider the syncretism of the acc and the gen of sound plurals. See the tableau in (15). I assume that an input, for example, comprises the lexeme MUʔMIN whose lexical stem is *muʔmin* and the feature value set {acc, pl}. I also assume that the relevant base is *-i* {gen}. The word *muʔmin-i* is the winning candidate although it violates the lower-ranked constraint {acc}: *-a*, which requires the accusative to be marked by the suffix *-a*. The illicit candidate **muʔmin-i* is ruled out by the grammar because it fatally violates the constraint {sound pl}: LV, which requires the sound plural to be marked by a long vowel. The illicit candidate **muʔmin-a* is also ruled out because it fatally violates both the constraints IDENT GA (VH) and {sound pl}: LV in that /a/ is a short vowel and does not correspond to /i/ of the genitive base.

(15) MUʔMIN (Classical Arabic)

MUʔMIN, acc, pl <i>muʔmin</i> Base: gen: <i>-i</i>	IDENT GA (VH)	IDENT AG (VH)	{sound pl}: LV	{acc}: <i>-a</i>	{gen}: <i>-i</i>
☞ a. MUʔMIN acc pl <i>muʔmin</i> <i>-i:</i>				*	
b. MUʔMIN acc pl <i>muʔmin</i> <i>-i</i>			*!	*	
c. MUʔMIN acc pl <i>muʔmin</i> <i>-a</i>	*!		*		

The same grammar can account for the syncretism of the acc and gen of diptotes. Consider the following tableau. Assume an input, for example, comprises the lexeme ʔASWAD whose lexical stem is *ʔaswad* and the feature value {gen}. I also assume that the relevant base is *-a* {acc}. Candidate (b) is ruled out because /i/ does not correspond to /a/ of the acc base.

(16) ʔASWAD (Classical Arabic)

ʔASWAD, gen <i>ʔaswad</i> Base: acc: <i>-a</i>	IDENT GA (VH)	IDENT AG (VH)	{sound pl}: LV	{acc}: <i>-a</i>	{gen}: <i>-i</i>
☞ a. ʔASWAD, gen <i>ʔaswad</i> <i>-a</i>					*
b. ʔASWAD, gen <i>ʔaswad</i> <i>-i</i>		*!			

To account for the syncretism of the abs(olutive) and acc(usative) of both male personal names (Type V nouns) and singular nouns (Type I nouns) in Diyari declension, I propose two crucial OO correspondence constraints and two realization constraints in (17).

- (17) a. MAX (acc (base), abs / male personal name): Every segment in the base accusative form has a correspondent in the form of the absolutive in the environment of a male personal name (mpn). (MAX (acc, abs))

- b. DEP (abs (base), acc / singular noun): Every segment in the form of the accusative has a correspondent in the base absolutive form in the environment of a singular noun. (DEP (abs, acc))
- c. {abs}: -∅: The absolutive is marked by a zero suffix in the output.
- d. {acc}: -*ḥa*: The accusative is marked by the suffix -*ḥa* in the output.

Let us first consider the syncretism of the accusative and absolutive of male personal names. Assume the input comprises a male personal name and the absolutive feature value and the relevant base is -*ḥa* {acc}. Consider the tableau in (18). The affix -*ḥa* is the winning candidate although it violates the lower-ranked constraint {abs}: -∅. The output candidate *-∅ fatally violates the OO correspondence constraint MAX (acc, abs) because the base form has no correspondent in the output.

(18) Male personal name (Diyari)

abs / mpn Base: acc: - <i>ḥa</i>	MAX (acc, abs)	DEP (abs, acc)	{abs}: -∅	{acc}: - <i>ḥa</i>
☞ a. abs / mpn - <i>ḥa</i>			*	
b. abs / mpn -∅	<i>ḥa</i>			

The same grammar can account for the syncretism of the accusative and absolutive of singular nouns. Assume the input contains a singular noun and the accusative feature value and the relevant base is -∅ {abs}. The output candidate *-*ḥa* fatally violates the OO correspondence constraint DEP (abs, acc) because the output has no correspondent in the base. See the tableau in (19).

(19) Singular noun (Diyari)

acc / singular Base: abs: -∅	MAX (acc, abs)	DEP (abs, acc)	{abs}: -∅	{acc}: - <i>ḥa</i>
☞ a. acc / singular -∅				*
b. acc / singular - <i>ḥa</i>		<i>ḥa</i>		

The constraint-based grammar in which OO correspondence constraints outrank realization constraints performs as well as rules of referral (Zwicky 1985, Stump 1993) in accounting for DBS. Take the syncretism of the nom sg and acc sg of default neuters in the Latin second declension as an example. Consider the rules in (20) (Baerman 2004: 816). The rule of referral in (20a) says that in the environment of a default neuter, the nom sg refers to the phonological form of the acc sg. This rule of referral feeds the rule of exponence in (20b) which spells out the acc sg. Unlike the impoverishment-plus-insertion theory, which makes an excessively restrictive prediction about directions of syncretism, the constraint-based grammar and rules of referral and exponence have no problem accounting for DBS.

- (20) a. nom sg in default neuter = acc sg
 b. acc sg = stem + *-um*

Additionally, the constraint-based grammar more clearly shows that cases of syncretism in which a direction has to be specified involve a *copying process*. The constraint-based grammar also captures the two functions of a rule of exponence, i.e., a rule of exponence not only spells out abstract morphosyntactic or semantic feature values but also sometimes provides a base which is to be copied by a distinct set of morphosyntactic feature values in a rule of referral.

A related question arises. What can act as a base which is to be copied when we need to specify the direction of syncretism? There are several possibilities which by no means form an exhaustive list. First, the phonological form realizing a feature value which occupies more paradigmatic cells of this feature value may act as a base (e.g., in the Latin second declension, *-um*, which is the phonological form realizing the accusative singular, acts as a base for the nominative singular of a default neuter to copy because *-um* occupies more accusative cells than nominative ones). Second, the phonological form which realizes a universally less marked feature value tends to be a base (Noyer 1998, Bobaljik 2001). Additionally, Albright (in press) argues that the plural form in early Yiddish was a base on which morphological leveling took place although the plural is a marked feature value. He remarks that “in this case, the plural is the form that most clearly exhibits lexical contrasts, and extending the plural variant does the least violence to recoverability.”

5.3 Partial Identity of Forms

In this section I show that output-to-output correspondence constraints are more fine-grained mechanisms than rules of referral although both can capture directional syncretism. OO correspondence constraints can also capture forms which are partially identical while by contrast rules of referral connect two fully identical forms.

5.3.1 Pinker 1998

Pinker 1998 observes that English words such as *workman* and *snowman* have the irregular inflection *X-men* while *Walkman* ‘a personal stereo’ doesn’t. Based on the

Right-hand Head Rule (Williams 1981), Pinker argues that the plural form of *Walkman* is *Walkmans* instead of **Walkmen* because something (let's say X) prevents *Walkman* from inheriting its manner of inflection from its rightmost morpheme *-man*. Pinker assumes the structure for *Walkman* is $[_N [_V \textit{Walk}] [_X [_N \textit{man}]]]$. Pinker's account leaves two questions unaddressed: (i) It is not clear what this "something" or X refers to. (ii) It is not clear why this X stands in between the nominal categories in cases like *Walkman*.

5.3.2 A realization OT account of partial identity of forms

I show that the ranking schema under which OO correspondence constraints outrank realization constraints can account for the distinction between *snowmen* and *Walkmans*. I organize nouns including *workman* and *snowman* with both the morpheme *-man* (/mæn/) and the sense of "human appearance" into one inflectional class (Aronoff 1994) in that they decline in the same way to realize the plural feature value. Let us call this class "*man-class*." I propose a relevant OO correspondence constraint and a realization constraint in (21).

(21) a. IDENT (pl / MAN), (pl / {N, *man-class*}) (F): Corresponding segments of both the plural exponent of MAN and a *man-class* noun have identical values for any phonological feature. (IDENT (pl / MAN), (pl / *man-class*))

b. {pl}: -s: Plural is realized by the suffix *-s* in the output.

Let us first consider *snowmen*. I assume that the input comprises the lexeme SNOWMAN and its stem *snowman* and the plural feature value. The relevant base is *men* which realizes both MAN and the plural feature value. Consider the tableau in (22). The word *snowmen* is the winning candidate although it violates the lower-ranked constraint {pl}: -s. The word *snowmen* satisfies IDENT (pl / MAN), (pl / *man-class*) because corresponding segments of both *men* and *snowmen* have identical phonological feature values.¹¹ (For the moment, I put aside issues concerning stress and vowel reduction which call for higher-ranked phonological constraints.) The illicit form **snowmans* fatally violates the OO correspondence constraint because the plural exponent *men* in the base does not correspond to the plural exponent *-s* in the output with respect to phonological features. The illicit form **snowmens* violates *FEATURE SPLIT because the plural feature value is realized twice.

(22) SNOWMAN

[SNOWMAN, <i>man-class</i>], pl <i>snowman</i> Base: [MAN, <i>man-class</i>], pl <i>men</i>	IDENT (pl / MAN), (pl / <i>man-class</i>)	*FEATURE SPLIT	{pl}: -s
a. [SNOWMAN, <i>man-class</i>], pl <i>snowmen</i>			*
b. [SNOWMAN, <i>man-class</i>], pl <i>snowman</i> -s	*!		
c. [SNOWMAN, <i>man-class</i>], pl <i>snowmen</i> -s		*!	

Next, let us consider *Walkmans*. I assume that the input comprises the lexeme WALKMAN and its stem *Walkman* and the plural feature value. The base is *men* which realizes both MAN and the plural feature value.¹² The constraint IDENT (pl / MAN), (pl / *man-class*) does not apply to *Walkmen* or *Walkmans*, because WALKMAN is not a *man-class* noun since WALKMAN does not denote the sense of “human appearance.” *Walkmen* is ruled out by the constraint {pl}: -s. *Walkmans* is the winning candidate which satisfies both the constraints IDENT (pl / MAN), (pl / *man-class*) and {pl}: -s. See the following tableau.

(23) WALKMAN

WALKMAN, pl <i>Walkman</i> Base: [MAN, <i>man-class</i>], pl <i>men</i>	IDENT (pl / MAN), (pl / <i>man-class</i>)	{pl}: -s
a. WALKMAN, pl <i>Walkman</i> -s		
b. WALKMAN, pl <i>Walkmen</i>		*!

This analysis captures the observation that the plural form of MAN is unpredictable¹³ while there is a productive process in which the plural forms of *man-class* nouns copy the plural form of MAN. Similar analyses apply to other inflectional classes in English such as the *go-class* including *go*, *forgo*, *undergo*, etc. and the *stand-class* including *stand*, *understand*, *withstand*, etc.

Additionally, this approach circumvents the problems for Pinker’s (1998) analysis of *Walkmans*. It straightforwardly shows that the plural form of WALKMAN does not

copy *men* because the meaning of the whole lexeme prevents WALKMAN from joining the *man*-class.

5.3.3 Rules of referral

OO correspondence constraints have a wider scope of application than rules of referral (Zwicky 1985, Stump 1993) which would encounter problems to account for partial identity of forms. In the spirit of Zwicky (1985), who uses rules of exponence to realize German suppletive determiners,¹⁴ we can propose a rule of exponence in (24) to realize the plural of the lexeme MAN.

(24) {MAN, pl} = *men*

We cannot, however, use a rule of referral like (25) to realize the plural form of the lexeme SNOWMAN because otherwise the plural form of SNOWMAN would be *men* instead of *snowmen*.¹⁵ Rule (25) says that the plural form of SNOWMAN is identical to the plural form of MAN which is *men*.

(25) {SNOWMAN, pl} = {MAN, pl}

5.4 Conclusions

I have shown that an OT approach based on both OO correspondence constraints and realization constraints can account for both directional syncretism and cases in which partially identical lexemes share the same inflectional formative. Divergent bidirectional syncretism which brings about both marked and unmarked forms poses a problem for the tenet of impoverishment-plus-insertion that the form of a less marked feature value always prevails. In contrast to both impoverishment-plus-insertion and referral, the constraint-based approach shows that directional syncretism involves a copying process in which the form of one set of morphosyntactic feature values copies that of the other. An approach based on the Right-hand Head Rule has problems accounting for nouns like WALKMAN and SNOWMAN which contain the same root but do not undergo the same inflectional process because it is not clear what prevents WALKMAN from being inflected in the same way as SNOWMAN. OO correspondence constraints are more fine-grained mechanisms and have a wider scope of application than rules of referral in that the former are able to account for both full and partial identity of forms while rules of referral connect two fully identical forms.

Notes

¹ See also Carstairs-McCarthy (1998), Baerman, Brown, and Corbett (2005) for criticisms of the impoverishment theory from a different perspective, i.e., if we reasonably manipulate the morphosyntactic feature values of vocabulary items, impoverishment will make different predictions about syncretic directions.

² Third declension neuter nouns like *tempus* ‘time’ pattern similarly to *vulgus*, in that the form of both the nom and acc resembles the masculine and feminine nominative (e.g., *dens* ‘tooth’, *miles* ‘soldier’).

³ I thank Robert Hoberman for explaining to me what diptotes are. “Mostly, it is those with certain morphophonological forms: nouns and adjectives of the shape ?aCCaC , plurals of the shape CaCaaCi(i)C , any noun or adjective (singular or plural) with the suffixes *-aa* or *-aa?*, and some personal names” (Bob Hoberman p.c.).

⁴ I thank Robert Hoberman for explaining to me the difference between a diptote and a triptote. “*?aswad* is basically a diptote. However, all diptotes revert to the default triptote inflection in certain environments, of which the most frequently occurring are (a) when preceded by the definite marker *al-* and (b) when they are the head of a construct phrase, which is a syntactic structure of this form: $[\text{DP N DP}]$. For example, the genitive case of ‘black’ is *?aswad-a*, but ‘the black (one) (GEN)’ is *al-?aswad-i*, ‘the black one (GEN) of the family’ is *?aswad-i al-?usraa*, and ‘the black one (GEN) of you (PL) is *?aswad-i=kum*” (Bob Hoberman p.c.).

⁵ In (3) the dative marker *-ō* seems to be the most suitable candidate for a default since it occupies the largest number of paradigmatic slots.

⁶ Apart from the feature identity constraints in (11), there are other OO correspondence constraints such as MAX-OO which bans the deletion of a segment of an output which has a correspondent in the base, and DEP-OO which bans the occurrence of a segment in the output which does not have a correspondent in the base. For the sake of brevity and simplicity I sometimes omit the discussion of some OO correspondence constraints which are not crucial to our results.

⁷ This constraint can also be formulated as follows. Corresponding segments of the forms of both the base acc sg and the nom sg plus a diacritic feature <default neuter> have identical values for any phonological feature.

⁸ Wunderlich (2000, 2005) describes syncretism on the basis of impoverishment and underspecification. Baerman (2004) criticizes Wunderlich’s (2005) treatment of syncretism and remarks that underspecification is incapable of accounting for DBS.

⁹ The notion of a base in this chapter is an output phonological form which is to be copied. By contrast, Kager (1999) gives a different definition of a base and proposes that a base should be a free-standing word and contains a subset of the grammatical features of the derived form.

¹⁰ The form **bell-um-us* (nom sg) is also an important candidate which satisfies both IDENT AN and {nom sg}: *-us*. It is, however, ruled out by the markedness constraint *FEATURE SPLIT which should rank higher than the two realization constraints in (12), because nominative is realized by both *-um* and *-us*.

¹¹ The word *snowmen* violates the constraint DEP-OO which requires no occurrence of additional segment compared to the base *men*. DEP-OO should therefore rank lower than MAX-IO which requires no deletion of the input segments of *snowman*. The output candidate **men* fatally violates MAX-IO and is therefore ruled out.

¹² It is possible that there is no base for WALKMAN to copy because by contrast all *man*-class nouns are free-standing words and have the semantic structure “something that looks like a man.” This assumption, however, encounters a problem when we account for, for example, the past tense forms of UNDERGO, FORGO, etc. which have *went* as their base. Verbs like UNDERGO, FORGO are semantically unrelated to GO.

¹³ If we follow Pinker 1998, we can assume that the irregular form *men* is listed in the lexicon. By contrast, Distributed Morphology (Halle and Marantz 1993) would assume that $-\emptyset$ marks the plural of the Root $\sqrt{\text{MAN}}$, which is followed by a readjustment rule triggering a Root-internal vowel change.

¹⁴ Zwicky’s rule of exponence is formalized as follows, for example. “[INDEX: 15, CASE: nom, GEND: neut, NUM: sg] is realized as /das/” (Zwicky 1985: 383).

¹⁵ The Head Application Principle (Stump 2001) accounts for the plural form of SNOWMAN, though it is not clear how it accounts for the plural form of WALKMAN.

Chapter Six

Conclusion

In this dissertation, I propose an inferential-realizational model of inflectional morphology (Matthews 1972, Zwicky 1985, Anderson 1992, Aronoff 1994, Stump 2001) within the framework of Optimality Theory (Prince and Smolensky 1993). Following Russell 1995, Yip 1998, Hyman 2003, MacBride 2004, I assume that the phonological information of inflectional affixes is introduced through realization constraints (RC) which associate abstract morphosyntactic or semantic feature values with phonological forms. I propose that rankings of realization constraints conform to the specificity condition, i.e. a constraint realizing a more specific morphosyntactic feature value set outranks a less specific realization constraint. I also propose that the unmarked situation in which one feature value is realized by one form (Wurzel 1989) is encoded in two universal and violable markedness constraints, *FEATURE SPLIT which bans the realization of a feature value by more than one form and *FEATURE FUSION which bans a form realizing more than one feature value.

Based on this model, I examine language phenomena such as OCP-triggered selection of phonologically unrelated (allo)morphs in Greek, Hungarian, Tswana, and Spanish, ordering of inflectional affixes in Lezgian, blocking of inflectional affixes and extended morphological exponence in languages like Tamazight Berber, and directional syncretism in languages like Latin.

In **Chapter 2**, I discuss both haplology and selection of phonologically unrelated (allo)morphs under the condition that the morphs in question are adjacent. I compare the proposed realization OT approach to morphologically restricted OT approaches which assume that all phonological information is introduced through inputs and morphological information such as “affix”, “root”, and “stem” constitutes enough input and output morphological information for the grammar to process. I show that although both OT models seem to be able to account for haplology, the realization OT approach readily accounts for OCP-triggered selection of phonologically unrelated (allo)morphs in languages such as Greek, Hungarian, Spanish, and Tswana. Cases of OCP-triggered (allo)morph selection in these languages involve the emergence of a morphosyntactically less specific form. I propose the ranking $OCP \gg RC_{specific} \gg RC_{less\ specific}$ and show that this ranking not only follows the specificity condition but also reflects the emergence of the unmarked, i.e., a morphosyntactically less specific (allo)morph emerges to avoid a violation of OCP constraints though it generally does not show up in the context where a morphosyntactically more specific (allo)morph is expected to occur. By contrast, a morphologically restricted OT approach needs to stipulate in the input the relation between a morphosyntactically more specific (allo)morph and a morphosyntactically less specific one (e.g. Bonet 2004). It also

needs to stipulate different input morphs in different morphosyntactic contexts. Such a stipulation will be avoided if we build certain morphemes into constraints. Morphosyntactic feature values are also necessary in accounting for Swedish haplology which varies with respect to meanings of affixes. An approach based on the generation of morphosyntactic feature values (Grimshaw 1997b, 2001, Wunderlich 2001) is incapable of accounting for OCP-triggered haplology or (allo)morph selection whose nature is morphophonological, because such an approach lacks a systematic mechanism to spell out morphosyntactic feature values. Finally, I discuss problems caused by voicing assimilation for realization approaches and show that there are several potential solutions though all of them have defects, which shows that the relation between realization (morphology) and phonological processes awaits further research.

In **Chapter 3**, I study Lezgian inflectional morphology with a focus on affix order. It is found that case markers are outside number markers. Locative markers which scope over localization markers are farther away from the nominal stem. Past tense markers are outside tense-aspect markers. Participles which express relative clauses are outside temporal-aspectual affixes. The negative marker in the indicative environment occurs between past tense and tense-aspect markers. The past tense suffix *-ir* does not appear in an affirmative context or follow tense-aspect markers, which always end in a vowel, because otherwise we would observe vowel hiatus, which is generally banned in Lezgian. I show that these generalizations are captured by both universal scopal and phonological constraints and language-particular realization and templatic constraints. The proposed realization OT approach readily captures universal generalizations on affix order (Greenberg 1963, Bybee 1985), e.g. a number exponent cannot be farther away from a nominal stem than a case exponent because case scopes over number. Such generalizations are missed in Paradigm Function Morphology (Stump 2001) without extraordinary machinery.

In **Chapter 4**, I show that the proposed realization OT model provides a unified approach to both blocking and extended exponence without recourse to either a distinction between primary and secondary exponents (Noyer 1992, 1997) or multiple rule blocks (Stump 2001). I focus my discussion on Tamazight Berber and Classical Arabic conjugations which have been widely discussed in the literature. I show that by ranking *FEATURE SPLIT lower than constraints realizing the same morphosyntactic or semantic feature values we observe extended exponence; otherwise, we observe blocking. The markedness constraint *FEATURE SPLIT is especially good at accounting for cases of discontinuous bleeding (Noyer 1992, 1997) in which blocking takes place across different positions. Compared to the proposed realization OT model, an OT model in which affixes are introduced through inputs needs to make bald stipulations to deal with both blocking and extended exponence. Incremental morphological models (e.g. Steele 1995, Wunderlich 1996) in which affixes introduce morphosyntactic or semantic feature values must be energetic in excluding extended exponence and especially cases in which a morphosyntactic or semantic feature value is realized by several exponents whose feature content is the same or among whose feature value sets there exists a subset relation because incremental models assume

that affixation is strictly information-adding.

In **Chapter 5**, I show that the proposed realization OT approach based on both output-to-output correspondence constraints and realization constraints can account for both directional syncretism and cases in which partially identical lexemes share the same inflectional formative. Divergent bidirectional syncretism which brings about both marked and unmarked forms poses a problem for the tenet of the feature impoverishment-plus-feature insertion theory (Noyer 1997, 1998) that the form of a less marked feature value always prevails. In contrast to both referral and impoverishment-plus-insertion, the realization OT approach shows that directional syncretism involves a copying process in which the form of one set of morphosyntactic or semantic feature values copies that of the other. Additionally, output-to-output correspondence constraints are more fine-grained mechanisms and have a wider scope of application than rules of referral in that the former are able to account for both full and partial identity of forms while rules of referral connect two fully identical forms.

To conclude, I show that the proposed realization OT model has advantages over other morphological models in several ways. It readily accounts for OCP-triggered selection of phonologically unrelated (allo)morphs and provides a unified approach to both blocking and extended exponence. By contrast, a morphologically restricted OT approach would miss a unified grammar and rely on bald stipulations to spell out a competing exponent in the input. Additionally, blocking and extended exponence need to be stipulated and rendered completely unrelated under a morphologically restricted OT approach. Without extraordinary machinery, an OT model based on the generation of morphosyntactic or semantic feature values is unable to capture morphological haplology, (allo)morph selection, universal generalizations on affix order, blocking and extended exponence, because such a model lacks a systematic non-cyclic mechanism inherent in OT to spell out abstract morphosyntactic or semantic feature values. Moreover, it is hard for ruled-based realizational models such as Paradigm Function Morphology and Distributed Morphology to incorporate universal generalizations on phonotactics and provide a unified approach to both blocking and extended exponence, which reflect the violability of *FEATURE SPLIT. Universal generalizations on affix order are missed in Paradigm Function Morphology without extraordinary machinery. Incremental morphological models which assumes that affixation is strictly information-adding need to exclude extended morphological exponence which occurs cross-linguistically and especially cases in which there exists a subset relation among the morphosyntactic or semantic feature value sets introduced by co-occurring affixal exponents. By contrast, extended exponence is fully compatible with the proposed realization OT model. Compared to mechanisms such as impoverishment-plus-insertion and referral in rule-based realizational models, the proposed realization OT approach can incorporate output-to-output correspondence constraints to capture both directional syncretism, which may bring about both marked and unmarked forms, and cases in which partially identical lexemes share the same inflectional formative.

Our realization OT model also makes predictions about impossible language phenomena. Assume that in a language L, the exponent *X*, which realizes a feature set subsuming feature values {a, b}, exceptionlessly blocks *Y*, which realizes a feature set subsuming feature values {b, c}. Additionally, assume that in L, *Y* exceptionlessly blocks *Z*, which realizes a feature set subsuming {a, c}. Then it is predicted that *X* would also exceptionlessly block *Z* because of the following ranking schema.

(1) *FEATURE SPLIT, {a, b, ...}: *X* >> {b, c, ...}: *Y* >> {a, c, ...}: *Z*

The exponent *X* would not co-occur with *Z* because otherwise {a} would be realized by both *X* and *Z*, which violates *FEATURE SPLIT.

References

- Abdel-Massih, Ernest T. 1971. *A reference grammar of Tamazight*. Center for Near Eastern and North African Studies. University of Michigan, Ann Arbor.
- Abraham, R. C. 1940. *The principles of Tiv*. Farnborough, England: Gregg.
- Albright, Adam. In press. Inflectional paradigms have bases too: evidence from Yiddish. In Asaf Bachrach and Andrew Nevins (eds.) *The Bases of Inflectional Identity*. Oxford University Press.
- Anderson, Stephen R. 1986. Disjunctive ordering in inflectional morphology. *Natural language and linguistic theory* 4: 1-31.
- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 2001. On some issues in morphological exponence. *Yearbook of Morphology* 2000: 1-18.
- Aronoff, Mark. 1976. *Word-formation in generative grammar*. Cambridge, Massachusetts: MIT Press.
- Aronoff, Mark. 1992. *Morphology by itself: Stems and inflectional classes*. Cambridge, Massachusetts: MIT Press.
- Austin, Peter. 1981. *A grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.
- Baerman, Matthew. 2004. Directionality and (un)natural classes in syncretism. *Language* 80:807-827.
- Baerman, Matthew, Dunstan Brown, and Greville G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.
- Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373-415.
- Beard, Robert. 1995. *Lexeme-Morpheme Base Morphology: A general theory of inflection and word-formation*. State University of New York Press, Albany.
- Beckman, Jill N. 1998. *Positional faithfulness*. Doctoral dissertation. University of Massachusetts, Amherst. Distributed by GLSA, UMass, Amherst.

- Bentolila, Fernand. 1981. *Grammaire fonctionnelle d'un parler berbère*. Paris: Société d'Etudes Linguistiques et Anthropologiques de France.
- Benua, Laura. 1995. Identity effects in morphological truncation. In *Papers in Optimality Theory* (University of Massachusetts Occasional Papers 18), ed. by Jill N. Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk, 77-136. GLSA, UMass, Amherst.
- Bobaljik, Jonathan David. 2001. Syncretism without paradigms: Remarks on Williams 1981, 1994. *Yearbook of Morphology 2001*, 53-85.
- Bonet, Eulàlia. 1991. *Morphology after syntax: Pronominal clitics in Romance*. Doctoral Dissertation. MIT.
- Bonet, Eulàlia. 1995. Feature Structure of Romance clitics. *Natural language and linguistic theory* 1995: 607-647.
- Bonet, Eulàlia. 2004. Morph insertion and allomorphy in Optimality Theory. In Paul Boersma and Juan Antonio Cutillas (eds.), *Advances in Optimality Theory*, special issue of *International Journal of English Studies* 4: 74-104.
- Börjars, Kersti. 2003. Morphological status and (de)grammaticalization: The Swedish possessive. *Nordic Journal of Linguistics* 26: 133-163.
- Brentari, Diane. 1998. Comments on the paper by Yip. In Steven G. Lapointe, Diane K. Brentari, and Patrick M. Farrell (eds.) *Morphology and its relation to phonology and syntax*. 247-258. Stanford, California: CSLI Publications.
- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Carstairs, Andrew. 1988. Some implications of phonologically conditioned suppletion. *Yearbook of Morphology* 1988: 67-94.
- Carstairs, Andrew. 1990. Phonologically conditioned suppletion. In Dressler, W. U., Luschützky, H. C., Pfeiffer, O. E. and Rennison, J. R. (eds.), *Contemporary Morphology*. 17-23. Berlin: Mouton de Gruyter.
- Carstairs-McCarthy, Andrew. 1998. Comments on the paper by Noyer. In *Morphology and its relation to phonology and syntax*, ed. by Steven G. Lapointe, Diane K. Brentari, and Patrick M. Farrell, 286-301. Stanford, California: CSLI Publications.

- Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Clahsen, Harald. 1999. Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences* 22: 991-1013.
- Cole, Desmond T. 1955. *An introduction to Tswana grammar*. Longmans and Green.
- Comrie, Bernard. 1975. Causatives and universal grammar. *Transactions of the Philological Society* (1974): 1-32.
- Comrie, Bernard. 1976. The syntax of causative constructions: Cross-language similarities and divergences. In *The grammar of causative constructions* (Syntax and Semantics 6), ed. by Masayoshi Shibatani, 261-312. New York: Academic Press.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbin, Danielle and Marc Plénat. 1992. Note Sur L'Haplologie des Mots Construits. *Langue Française* 96: 101-112.
- de Lacy, Paul. 1999. Haplology and correspondence. In Paul de Lacy and Anita Nowak (eds.) *Papers from the 25th Anniversary*. University of Massachusetts Occasional Papers (UMOP) 24. Amherst, MA: GLSA Publications, 51-88. [Also Rutgers Optimality Archive #298.]
- Dressler, Wolfgang U. 1977. Phono-morphological dissimilation. In Wolfgang U. Dressler and Oskar E. Pfeiffer (eds.) *Phonologica 1976*: 41-48. Innsbruck.
- Embick, David, and Morris Halle. 2005. On the status of stems in morphological theory. In T. Geerts and H. Jacobs (eds.) *Proceedings of Going Romance 2003*. John Benjamins.
- Embick, David and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32: 555-595.
- Fischer, Wolfdietrich. 1997. Classical Arabic. In *The Semitic languages*, ed. by Robert Hetzron, 187-219. London: Routledge.
- Goldsmith, John A. 1976. *Autosegmental phonology*. Ph.D dissertation. MIT. Published 1979, New York: Garland.
- Goldsmith, John A. 1984. Meeussen's Rule. In Mark Aronoff and Richard T. Oehrle (eds.) *Language sound structure*. 245-259. MIT Press.

- Golla, Victor. 1970. Hupa Grammar. Ph.D dissertation. University of California at Berkeley.
- Golston, Chris. 1995. Syntax outranks phonology: Evidence from Ancient Greek. *Phonology* 12:343-368.
- Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph Greenberg (ed.) *Universals of language*. Cambridge, MA: MIT Press.
- Grimshaw, Jane. 1986. A morphosyntactic explanation for the Mirror Principle. *Linguistic Inquiry* 17: 745-749.
- Grimshaw, Jane. 1997a. Projection, heads, and Optimality. *Linguistic Inquiry* 28: 373-422.
- Grimshaw, Jane. 1997b. The best clitic: Constraint conflict in morphosyntax. In Liliane Haegeman (ed.) *Elements of grammar*. Kluwer Academic Publishers.
- Grimshaw, Jane. 2001. Optimal clitic positions and the Lexicon in Romance clitic systems. In G. Legendre, J. Grimshaw, and S. Vikner eds. *Optimality-Theoretic Syntax*. MIT Press.
- Halle, Morris. 1997. Distributed Morphology: Impoverishment and Fission. *MIT Working Papers in Linguistics* 30: 425-449. Also in J. Lecarme, J. Lowenstein, U. Shlonsky (eds.) 2000. *Research in Afroasiatic Grammar: Papers from the Third Conference on Afroasiatic Languages, Sophia Antipolis France, 1996*. 125-151. Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111-176. Cambridge, MA: MIT Press.
- Harris, Alice C. 2007. Exuberant exponence in Tsova-Tush. Ms. Stony Brook University.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin, New York: Mouton de Gruyter.
- Hayes, Bruce. 1986. Inalterability in CV phonology. *Language* 62: 321-352.
- Hooper, Joan. 1976. *Introduction to Natural Generative Phonology*. New York: Academic Press.

- Hyman, Larry M. 2003. Suffix ordering in Bantu: A morphocentric approach. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2002*. Dordrecht: Kluwer Academic, 245-281.
- Julien, Marit. 2000. Syntactic heads and word-formation: A study of verbal inflection. Ph.D dissertation, University of Tromsø.
- Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Cambridge, MA: Blackwell Publishers Inc.
- Kenstowicz, Michael. 1997. Base identity and uniform exponence: alternatives to cyclicity. In J. Durand and B. Laks (eds.) *Current trends in phonology: Models and methods*, 363-394. Salford: University of Salford.
- Kibrik, A. E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka, II: Taksonomičeskaja grammatika*. (Publikacii otdelenija strukturnoj i prikladnoj lingvistiki, 12). Moscow: Izdatel'stvo Moskovskogo universiteta.
- Kiparsky, Paul. 1973. "The Elsewhere Condition." In Stephen R. Anderson & Paul Kiparsky (eds.) *A Festschrift for Morris Halle*, 93-106. New York: Holt, Rinehart and Winston.
- Kiparsky, Paul. 2005. Blocking and periphrasis in inflectional paradigms. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Springer, 113-135.
- Kurusu, Kazutaka. 2000. Double morphemic exponence as morphological opacity. Chicago Linguistic Society: 163-177.
- Lawrence, Wayne P. 1997. Haplology and vowel underspecification. In *Report of the special research project for the typological investigation of the languages and cultures of the East and West*. 381-388.
- Leben, William. 1973. *Suprasegmental phonology*. Ph.D dissertation. MIT. Published, New York: Garland.
- Leben, William. 1978. The representation of tone. In Victoria Fromkin (ed.) *Tone: A linguistic survey*. 177-219. New York: Academic Press.
- Li, Charles N. and Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.

- Lieber, Rochelle. 1992. *Deconstructing morphology*. University of Chicago Press.
- Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural language and linguistic theory* 17: 267-302.
- Lombardi, Linda. 2004. Positional faithfulness and voicing assimilation in Optimality Theory. In John J. McCarthy (ed.) *Optimality Theory in phonology: A reader*. 343-364. Blackwell Publishing Ltd.
- Lumsden, John. 1992. Underspecification in grammatical and natural gender. *Linguistic Inquiry* 23.3: 469-486.
- MacBride, Alexander Ian. 2004. A constraint-based approach to morphology. Doctoral dissertation, UCLA.
- Magometov, A. A. 1965. *Tabasaranskij jazyk*. Tbilisi: Mecniereba.
- Magometov, A. A. 1970. *Agul'skij jazyk*. Tbilisi: Mecniereba.
- Maiden, Martin. 2000. Phonological dissimilation and clitic morphology in Italo-Romance. In Lori Repetti (ed.) *Phonological theory and the dialects of Italy*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Matthews, P. H. 1972. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.
- Matthews, P. H. 1991. *Morphology*. Cambridge: Cambridge University Press.
- McCarthy, John J. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In Mercè González (ed.) *Proceedings of the Northeast Linguistic Society* 24: 333-379. GLSA, Amherst, MA.
- McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* 16: 331-399.
- McCarthy, John J. 2005. Optimal paradigms. In Laura J. Downing, T. A. Hall, and Renate Raffelsiefen (eds.) *Paradigms in phonological theory*. Oxford and New York: Oxford University Press.
- McCarthy, John J. and Alan Prince. 1993a. Generalized alignment. *Yearbook of Morphology* 1993: 79-153.
- McCarthy, John J. and Alan Prince. 1993b. Prosodic Morphology I: Constraint interaction and satisfaction. Ms. University of Massachusetts, Amherst, and Rutgers University.

- McCarthy, John J. and Alan Prince. 1995. Faithfulness and reduplicative identity. In *Papers in Optimality Theory* (University of Massachusetts Occasional Papers 18), ed. by Jill N. Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk, 249-384. GLSA, UMass, Amherst.
- McCarthy, John J. and Alan Prince. 2004. Generalized alignment: The prosody-morphology interface. In John J. McCarthy (ed.) *Optimality Theory in phonology: A reader*. 451-463. Blackwell Publishing Ltd.
- Menn, Lise and Brian MacWhinney. 1984. The repeated morph constraint: Toward an explanation. *Language* 60: 519-541.
- Mohanan, K. P. and Tara Mohanan. 2003. Universal and language-particular constraints in OT-LFG. In Miriam Butt and Tracy Holloway King (eds.) *Proceedings of the LFG03 Conference*. CSLI Publication.
- Moor, Marianne. 1985. *Studien zum lesgischen Verb*. Wiesbaden: Otto Harrassowitz.
- Neeleman, Ad and Hans van de Koot. 2005. Syntactic OCP Effects. In M. Everaert and H. van Riemsdijk (eds.) *The Blackwell Syntactic Companion*. London: Blackwell.
- Nicklas, Thurston D. 1972. The elements of Choctaw. University of Michigan dissertation.
- Norde, Muriel. 1997. *The history of the genitive in Swedish*. Dissertation. Vakgroep Skandinavische taal-en letterkunde, University of Amsterdam.
- Norde, Muriel. 2001a. Deflexion as a counterfactual factor in grammatical change. *Language Sciences* 23: 231-264.
- Norde, Muriel. 2001b. The history of the genitive: The full story. *Nordic Journal of Linguistics* 24: 107-118.
- Noyer, Rolf. 1992. Features, positions, and affixes in autonomous morphological structure. Doctoral dissertation. MIT.
- Noyer, Rolf. 1997. *Features, positions, and affixes in autonomous morphological structure*. New York & London: Garland Publishing, Inc.
- Noyer, Rolf. 1998. Impoverishment theory and morphosyntactic markedness. In *Morphology and its relation to phonology and syntax*, ed. by Steven G. Lapointe, Diane K. Brentari, and Patrick M. Farrell, 264-285. Stanford, California: CSLI Publications.

- Ortmann, Albert and Alexandra Popescu. 2000. Haplology involving morphological bound and free elements: Evidence from Romanian. *Yearbook of Morphology* 2000: 43-70.
- Paster, Mary. 2005. Pulaar verbal extensions and phonologically driven affix order. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2005*. Dordrecht: Springer, 155-199.
- Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. Holt, Rinehart and Winston, Inc.
- Pinker, Steven. 1998. Words and rules. *Lingua* 106:219-242.
- Plag, Ingo. 1998. Morphological haplology in a constraint-based morpho-phonology. In Wolfgang Kehrein and Richard Wiese (eds.) *Phonology and morphology of the Germanic languages*. 199-215. Tübingen: Niemeyer.
- Prince, Alan. 1984. Phonology with tiers. In Mark Aronoff and Richard T. Oehrle (eds.) *Language sound structure*. 234-244. MIT Press.
- Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms. Rutgers University, New Brunswick, and University of Colorado, Boulder.
- Pulleyblank, Douglas. 1986. *Tone in Lexical Phonology*. Dordrecht: Reidel.
- Radford, Andrew. 1977. Counter-filtering rules. *York Papers in Linguistics* 7.7-45.
- Rice, Keren. 2000. *Morpheme order and semantic scope: Word formation in the Athapaskan verb*. Cambridge: Cambridge University Press.
- Rosenthal, Sam. 1997. The distribution of prevocalic Vowels. *Natural Language and Linguistic Theory*, 15: 139-180.
- Ross, John Robert. 1972. Doubl-ing. *Linguistic Inquiry* 3: 61-86.
- Rounds, Carol. 2001. *Hungarian: An essential grammar*. London and New York: Routledge.
- Russell, Kevin. 1995. Morphemes and candidates in Optimality Theory. Ms. Rutgers Optimality Archive 44-0195.

- Russell, Kevin. 1997. Optimality Theory and morphology. In Diane Archangeli and D. Terence Langendoen (eds.) *Optimality Theory: An overview*. 102-133. Blackwell Publishers Inc.
- Schein, Barry and Donca Steriade. 1986. On geminates. *Linguistic Inquiry* 17: 691-744.
- Siptár, Péter and Miklós Törkenczy. 2000. *The phonology of Hungarian*. Oxford: Oxford University Press.
- Spencer, Andrew. 2003. Putting some order into morphology: Reflections on Rice (2000) and Stump (2001). *Journal of Linguistics* 39: 621-646.
- Steele, Susan. 1995. Towards a theory of morphological information. *Language* 71: 260-309.
- Stemberger, Joseph. 1981. Morphological haplology. *Language* 57.4: 791-817.
- Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Ph.D dissertation. MIT.
- Steriade, Donca. 2005. Chained correspondence and Romanian agentive expressions. Handout delivered at Stony Brook University.
- Stump, Gregory T. 1993. On rules of referral. *Language* 69: 449-479.
- Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Trommer, Jochen. 2003. The interaction of morphology and syntax in affix order. *Yearbook of Morphology* 2002: 283-324.
- Uslar, Petr Karlovič. 1896. *Ètnografija Kavkaza*. Jazykoznanie. VI. Kjurinskij jazyk. [Ethnography of the Caucasus. Linguistics. VI. The Küre language.] Tiflis.
- Vago, Robert M. 1980. *The sound pattern of Hungarian*. Washington, D. C.: Georgetown University Press.
- Wessén, Elias. 1965. *Svensk språkhistoria*. Stockholm: Almqvist and Wiksell.
- Williams, Edwin. 1981. On the notions “lexically related” and “head of a word”. *Linguistic inquiry* 12:245-274.
- Wilson, Stephen and Ayşe Pinar Saygin. 2001. Adverbs and functional heads in Turkish: Linear order and scope. *Proceedings of WECOL*.

- Woolford, Ellen. 2001. Case patterns. In *Optimality-Theoretic Syntax*, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner, 509-543. Cambridge, Massachusetts: MIT Press.
- Wunderlich, Dieter. 1996. Minimalist morphology: the role of paradigms. *Yearbook of Morphology 1995*: 93-114.
- Wunderlich, Dieter. 2001. A correspondence-theoretic analysis of Dalabon transitive paradigms. *Yearbook of Morphology 2000*, 233-252.
- Wunderlich, Dieter. 2005. Is there any need for the concept of directional syncretism? In *Explorations in nominal inflection*, ed. by Lutz Gunkel, Gereon Müller, and Gisela Zifonun. Berlin: Mouton de Gruyter, to appear.
- Wunderlich, Dieter and Ray Fabri. 1996. Minimalist Morphology: An approach to inflection. *Zeitschrift für Sprachwissenschaft 1995*, 14.2: 236-294.
- Wurzel, Wolfgang Ullrich. 1989. *Inflectional morphology and naturalness*. Dordrecht: Kluwer Academic Publishers.
- Yip, Moira. 1993. The interaction of ALIGN, PARSE-PLACE and ECHO in reduplication. Paper presented at the Rutgers Optimality Workshop.
- Yip, Moira. 1998. Identity avoidance in phonology and morphology. In Steven G. Lapointe, Diane K. Brentari, and Patrick M. Farrell (eds.) *Morphology and its relation to phonology and syntax*. 216-246. Stanford, California: CSLI Publications.
- Zwicky, Arnold. 1985. How to describe inflection. In *Proceedings of BLS 11*, ed. by Mary Niepokuj, Mary van Clay, Vassiliki Nikiforidou, and Deborah Feder, 372-386. University of California, Berkeley.