A Formal Analysis of Inflectional Marking in the Albanian Noun Phrase

Research Thesis

Presented in Partial Fulfillment of Graduation with Research Distinction in Linguistics in the Undergraduate Colleges of The Ohio State University

by

Kyle Maycock

The Ohio State University

2021

Thesis Committee:

Dr. Andrea Sims, Advisor

Dr. Brian Joseph

Dr. Robert Levine

Abstract

The Albanian noun phrase marks four morphosyntactic properties: number, gender, case, and definiteness. Every lexical word in the phrase mark number and gender, but only the first lexical word in the phrase—either a noun (1) or an adjective (2)—marks case and definiteness.

(1) vajz-at e mir-a
girl-NOM.DEF.PL.F NOM.DEF.PL.F good-PL.F
'the good girls'

(2) e mir-at vajz-a
NOM.DEF.PL.F good-NOM.DEF.PL.F girl-PL.F
'the good girls

Number and gender are straightforwardly morphological, but the placement of case and definiteness is dependent upon the syntax. In this way, this exponent is a clitic.

The Albanian clitic is especially informative about the morphology-syntax interface because of its "special" (Zwicky 1977) placement after the first lexical word, or second position (2P), and its cumulative exponence. There are many models of 2P clitic placement that treat 2P clitics as phrasal affixes, notably Halpern (1995) and Anderson (2005), but the Albanian clitic's cumulative exponence poses a problem for these models due to its noncanonical nature. In this thesis, I develop an analysis of the clitic using Head-Driven Phrase Structure Grammar (Pollard and Sag 1994) that accounts for the clitic as edge inflection, rather than treating it as phrasal affixation.

The clitic's cumulative exponence results in two paradigms for lexemes depending on their location within the phrase; when the word is in first position, it marks a larger set of properties than when it is in subsequent positions. This poses a problem to morphology, as it suggests morphology is privy to syntactic placement. In this thesis, I develop an analysis using Paradigm Function Morphology that allows morphology to remain blind to phrasal position.

Acknowledgements

This project was above and away the hardest project I've undertaken. There are many people without whom this project would never have been completed. Perhaps most importantly, I would like to thank my advisor, Dr. Andrea Sims, for the guidance and incredibly helpful comments that helped shape and embetter my thoughts and arguments for this thesis and any project that I undertake in the future. I would also like to thank my thesis defense committee, Dr. Robert Levine and Dr. Brian Joseph, for their encouragement and thorough comments. Rexhina Ndoci was also incredibly helpful by assisting in confirming Albanian facts, for which I am immensely grateful. And last, I would like to thank my support structure—my parents and fiancée—who endured my incoherent ramblings about this project and encouraged me anyway.

Table of Contents

Abbreviations

3	Third Person
ABL	Ablative
ACC	Accusative
AUX	Auxiliary
DAT	Dative
DEF	Definite
F	Feminine
GEN	Genitive
M	Masculine
NDEF	Indefinite
N	Neuter
NOM	Nominative
PST	Past tense
PL	Plural
SG	Singular

Chapter 1: Goals of this Thesis

1.1 Clitics and the Morphology-Syntax Interface

Albanian noun phrases mark four morphosyntactic properties: case, definiteness, number and gender. All lexical words in the phrase mark number and gender, but only the first lexical word—either the noun (1.1) or the adjective (1.2)—marks case and definiteness.

(1.1) vajz-at e mir-a
girl-NOM.DEF.PL.F NOM.DEF.PL.F good-PL.F
'the good girls'

(1.2) e mir-at vajz-a
NOM.DEF.PL.F good-NOM.DEF.PL.F girl-PL.F
'the good girls

Number and gender behave like morphological affixes, while the properties of case and definiteness behave like clitics. Clitics lie at the boundary of morphology and syntax, and are thus informative for study.

The Albanian clitic also exhibits "special" (Zwicky 1977) placement, being located after the first word, or in second position. Another such example is one from Serbian[1] (1.3).

(1.3) Čovek je voleo Mariju
man.NOM.SG AUX.3.SG love.PST.M Maria.ACC
'the man loved Maria'

The auxiliary particle, *je*, is in second position. The word order of this phrase is relatively free, providing the possible combinations seen in (1.4).

(1.4) Čovek je Mariju voleo
Voleo je čovek Mariju

[1] Data pulled from Halpern (1995), at the time referred to as Serbo-Croatian. Now, Serbian and Croatian are separate languages that are different enough to warrant referring to specifically Serbian.

Voleo je Mariju čovek
Mariju je čovek voleo
Mariju je voleo čovek

Regardless of what word is on the left edge of the phrase, *je* is always in second-position. This is similar to the Albanian clitic; regardless of the category of the word on the left edge of the phrase, the properties of case and definiteness attach in second-position. However, what further sets the Albanian data apart from the Serbian example above is its cumulative exponence. The two sets of properties have different behavior; the clitic properties (case and definiteness) are marked by only one word and the morphological properties (number and gender) are expressed by all words. Despite this differing behavior, they are expressed simultaneously by the *same* suffix. The cumulative expression seen in the Albanian noun phrase has not been previously analyzed and presents a problem for many prior models of clitic placement, such as Halpern (1995) and Anderson (2005).

The cumulative expression of the clitic also creates an interesting issue for morphology, as it creates varying levels of property expression. The first lexical word in the phrase marks *more* properties than subsequent words in the phrase. This presents an interesting issue in that lexemes have two different paradigms depending on location. Thus, an analysis of the data must deal with two major issues: placement in 2P, and the cumulative expression of case, definiteness, number, and gender.

1.2 Thesis Questions

In this thesis, the questions I ask are:

- How do we account for case and definiteness appearing in second position?
- How do we account for words having different forms depending on their position in the phrase (e.g. *vajzat* vs *vajza*)?
- What do the data tell us about the morphology-syntax interface?

The issue posed to syntax is the cumulative exponence of the suffix. The two sets of properties in the suffix have a disparity in placement domain; number and gender are clearly placed morphologically, whereas case and definiteness are placed with regard to the phrase, which is within the domain of syntax. As I show in chapter 3, many models (e.g., Halpern 1995 and Anderson 2005) contain a few key assumptions that struggle to deal with this phenomenon. Most importantly, they assume the insertion of form either simply occurs before placement, or that form licenses final placement. This assumption leaves us with multiple suffixes, when only one is licensed. As I also show in chapter 3, a model that does not make this assumption, such as Head-Driven Phrase Structure Grammar (Pollard and Sag 1994), is able to adequately explain the data.

While cumulative exponence poses a serious issue for the placement of the clitic, it poses no issue to morphology, assuming a model of morphology such as Paradigm Function Morphology (Stump 2001). The cumulative exponence of the first suffix causes it to have a different, more specific set of properties than all subsequent suffixes in the phrase, which may initially seem like an issue. However, using underspecification, as well as other various theoretical techniques, a Paradigm Function Morphology analysis, developed in chapter 4, adequately handles the data.

This thesis helps divide the roles of syntax and morphology regarding clitics into roles of placement and insertion of form, respectively. In this thesis, I argue against models that assume a more indistinguishable view of syntax and morphology, such as Distributed Morphology (Siddiqi 2019; Embick 2015). By using lexicalist models (Montermini 2019), or models that posit the primitives of syntax are words, like Head-Driven Phrase Structure Grammar, the placement of the properties is able to maintain the clitic's cumulative expression. This word is then relayed to

morphology, where “second position” is a result of suffixation. Models that do not function this way struggle with the data’s cumulative exponence.

1.3 Thesis Overview

This thesis is structured as follows. Chapter 2 introduces the Albanian noun morphology data and relevant theoretical concepts. Chapter 3 introduces the issue posed to syntax by the data’s cumulative exponence and showcases how the Albanian clitic behaves like edge inflection rather than the generally assumed phrasal affix category. Chapter 4 showcases how cumulative exponence poses no issue for a model of morphology like Paradigm Function Morphology. Chapter 5 concludes and summarizes the thesis with a discussion of how this thesis helps divide the roles of syntax and morphology at the morphology-syntax interface.

1.4 Description of Major Data Sources

This thesis pulls data from two major sources: *The Oxford English-Albanian Dictionary* (Newmark 1998) and *Standard Albanian: A Reference Grammar for Students* (Newmark et al. 1982). Nearly all examples in the data, especially in Chapter 2, are pulled from Newmark et al. (1982), which is the standard grammar for studying Albanian.

1.5 Terms and Concepts

Here, I define a few key terms and concepts. These concepts can have a wide range of meanings, especially when no clear definition is possible. In order to have as useful of a discussion as possible, I here define how these terms are used for the purpose of my argumentation.

1.5.1 Clitics and Second-Position

Clitics are a difficult phenomenon to define, and are notoriously difficult to delineate from affixes. The most notable attempt at clitic-affix delineation is Zwicky and Pullum (1983)'s criteria. These criteria are as follows:

A. "Clitics can exhibit a low degree of selection with respect to their hosts, while affixes exhibit a high degree of selection with respect to their stems." (p. 503)
B. "Arbitrary gaps in the set of combinations are more characteristic of affixed words than of clitic groups." (p.504)
C. "Morphophonological idiosyncrasies are more characteristic of affixed words than of clitic groups." (p. 504)
D. "Semantic idiosyncrasies are more characteristic of affixed words than of clitic groups." (p. 504)
E. "Syntactic rules can affect affixed words, but cannot affect clitic groups." (p. 504)
F. "Clitics can attach to material already containing clitics, but affixes cannot." (p. 504)

There are a number of known problems with the criteria (for discussion, see Sims and Joseph 2018 and Spencer and Luís 2012), but what is ultimately relevant here is that the purpose of this thesis is how the Albanian clitic challenges many previous models of clitic placement. To better exemplify how the Albanian clitic challenges these models in Chapter 3, I use Spencer and Luís (2012)'s canonical description. I use this description because I compare not the Albanian clitic to affixes, but the Albanian clitic to other clitics, namely the Serbian auxiliary particle and the English genitive exponent. By pinpointing canonical nature, I can better showcase the differences between these clitics.

Spencer and Luís (2012) pinpoint the canonical nature of clitics by comparing them to the canonical nature of affixes and words. They do so by outlining three major axes: content, form, and placement. Canonical affixes contain morphosyntactic content, rather than the lexical content of canonical words. Affixes also lack stress, in contrast to the stress bearing canonical word. Canonical words are placed with regard to the phrase, which means their placement is within the realm of syntax. This is in contrast to the placement of affixes, which is with regard to

a word, as in prefixation and suffixation on their stem. This placement is within the realm of morphology. Canonical clitics, then, fall somewhere between these. They share with affixes in morphosyntactic content and prosodically weak form, but share with words in their domain of placement: syntax.

Another important term regarding clitics is the idea of *phrasal affix* versus *edge inflection* (Spencer and Luís 2012). The key here is that a phrasal affix is a clitic that acts as an affix, but with its host as a phrase rather than the traditional word. The Serbian auxiliary in (1.3) is one such example. In contrast to phrasal affixes, edge inflection clitics, such as the English genitive exponent in (1.5) are less canonical in nature (the English data pulled from Spencer and Luís (2012) and adapted to American pronunciation).

(1.5) a. The girls' /gɚːlz/ names
b. *The girls's /gɚːlzəz/ names

Clitics categorized as edge inflection are still phrasally placed, but the important difference is that they display allomorphy with their host. A phrasal affix, in contrast, is more self-contained and is not expected to do so. In (1.5), allomorphy is displayed. If the genitive exponent were a self-contained unit like a phrasal affix, then we would expect the result in (1.5b). But haplology, a form of allomorphy, occurs here. One might consider this to be phonological rather than morphological, but (1.6) shows this to not be the case.

(1.6) The cheese's /t͡ʃizəz/ aroma

Since the haplology in (1.5a) is not present in (1.6), we can only assume this haplology is interaction with the host. The Albanian data also display a large amount of allomorphy, and so are categorized as edge inflection. It is this categorization that becomes relevant in chapter 3, since many previous models generally assume clitics are of the phrasal affix type. This

assumption poses a particular problem when applied to edge inflection clitics such as the Albanian clitic.

The placement of clitics is also an important concept, especially the concept of second-position (2P). A 2P clitic is one that falls immediately after the first item, although it is not entirely clear what that "anchoring" (Anderson 2005) element may be. Halpern (1995) notes two different types of this: the second word (2W) and the second daughter (2D). Clitics that fall after the *first word* are categorized as 2W, and clitics that fall after the *first phrase* are categorized as 2D. There is still some gray area here, however. The first word must often still be stress-bearing, which is a characteristic of canonical, lexical words. This means that morphosyntactic words that appear first are often not the host of a 2P clitic. This description is still a little off, as there are examples of the first word being stress bearing *and* morphosyntactic in content, but are still not the host of the clitic. For the purposes of this thesis, I define second position as after the first *lexical* word.

1.5.2 Cumulative Exponence

Cumulative exponence is a term that means the exponent in question is representing multiple morphosyntactic properties. This is particularly interesting for the clitic, since the first word in the phrase has a suffix cumulatively expressing both the clitic properties of case and definiteness and the non-clitic properties of number and gender. Because of this cumulative expression, the Albanian clitic presents a unique problem to previous models of syntax, as I show in chapter 3. This is also relevant in chapter 4, as the first lexical word in the phrase cumulatively expresses a larger amount of properties than the expression of subsequent words. This creates two different paradigms; one for the word expressing case and definiteness, and

another for when it is not. This is interesting from a morphological perspective, since we do not normally think of a lexeme as having two paradigms based on its location within the phrase.

Chapter 2: Introduction

In this chapter, I introduce the data and the problems presented by the clitic's cumulative expression. In 2.1, I introduce the inflectional morphology of the Albanian noun phrase, and in section 2.2 I introduce the relevant syntax of the Albanian noun phrase.

2.1 Albanian Noun Phrase Morphology

The scope of this thesis is the interaction between syntax and morphology within the Albanian noun phrase. Specifically, the scope is over the placement of the morphosyntactic properties of case and definiteness and the realization of the properties within the phrase. The interaction with phonology falls outside the bounds of this scope, including: phonologically conditioned phonological allomorphy, such as the insertion of *j* between vowels (*deleje*); and the deletion of ë (phonologically [ə]) in various locations, most notably next to other vowels (*vajzë*, *vajz-a*). Only phonologically conditioned allomorphy is ignored for the purposes of the thesis; all other allomorphy is accounted for.

In this section, I introduce the inflectional suffix paradigms, the allomorphy therein, and the difference between suffixes that attach to the first word in the phrase and and those that attach to subsequent words. I separate the singular and plural paradigms into separate sections because of their unusual characteristics; some words that are in separate singular classes are in the same plural class, and some that are in the same singular class are in separate plural classes. Section 2.1.1 and 2.1.2 will be centered on words that fall first in the phrase, and section 2.1.3 will be on subsequent words in the phrase. Section 2.1.4 focuses on adjectives and their differences from nouns.

2.1.1 The Singular Paradigms

The noun phrase in Albanian marks four distinct morphosyntactic properties: gender, number, case, and definiteness. There are five cases: nominative, accusative, dative, ablative, and genitive. For most of the paradigms, dative, ablative, and genitive are nearly identical. Ablative is morphologically differentiated only in the indefinite plural, and genitive is differentiated from the others with a requisite prephrasal particle (2.1). Otherwise, their suffixes are identical.

(2.1) lis-i i burr-it
oak-NOM.DEF.M.SG GEN man-GEN.DEF.M.SG
'the man's oak'

There are three genders in nouns: masculine, feminine, and neuter. Masculine and neuter, shown in their singular forms in Table 2.1, are similar in all paradigm cells except in the definite nominative and accusative.

		Masculine					Neuter	
		lis 'oak'	mik 'friend'	vëlla 'brother'	burrë 'man'	zë 'voice'	të folur 'speaking'	të ftohtë 'cold'
Indefinite	Nominative	lis	mik	vëlla	burrë	zë	të folur	të ftohtë
	Accusative	lis	mik	vëlla	burrë	zë	të folur	të ftohtë
	Dative	lis-i	mik-u	vëlla-i	burr-i	zër-i	të folur-i	të ftoht-i
	Ablative	lis-i	mik-u	vëlla-i	burr-i	zër-i	të folur-i	të ftoht-i
	Genitive	i lis-i	i mik-u	i vëlla-i	i burr-i	i zër-i	i të folur-i	i të ftoht-i
Definite	Nominative	lis-i	mik-u	vëlla-i	burr-i	zër-i	të foluri-t	të ftohtë-t
	Accusative	lis-in	mik-un	vëlla-në[2]	burr-in	zër-in	të foluri-t	të ftohtë-t
	Dative	lis-it	mik-ut	vëlla-it	burr-it	zër-it	të folur-it	të ftoht-it
	Ablative	lis-it	mik-ut	vëlla-it	burr-it	zër-it	të folur-it	të ftoht-it
	Genitive	i lis-it	i mik-ut	i vëlla-it	i burr-it	i zër-it	i të folur-it	i të ftoht-it

Table 2.1

Masculine and neuter singular suffixes begin with the underlying vowel /i/. This vowel will change to [u] (like *mik-u*) in a specific environment (Newmark et al. 1982: 159): following stems that end in *k*, *g*, *h*, or a stressed vowel (except *ó*). In general, when a stem ends with [ə]

[2] While the masculine definite accusative singular form is *-in*, there are some fossilized, archaic, and increasingly outdated examples of the form realizing as *-në* (Newmark et al. 1982: 160).

(orthographically *ë*), it will delete when neighboring another vowel. This is the case in instances like *burri*, instead of **burrëi*. Neuter definite nominative and accusative singular take the suffix *-t*, but, according to Newmark et al. (1982), stems that are derived from participles that end in *ur*, such as TË FOLUR, will insert *i* between the stem and the suffix (*të folurit)*. This explains why TË FTOHTË 'cold' takes *të ftohtët* instead of **të ftohtit*. There is a specific set of masculine nouns that end in *ë* that insert *r* before the suffix, which are usually monosyllabic with an absent coda. In all, there are a total of two singular classes here, one for morphosyntactically *masculine* nouns (Class I) and one for *neuter* nouns (Class II), although this does not line up with semantic gender. Similarly, there is one class (Class III) for *feminine* nouns.

Where masculine and neuter differ only slightly, the singular feminine paradigms (Table 2.2) are starkly different.

		vájzë 'girl'	rrufé 'thunderbolt'	déle 'sheep'
Indefinite	Nominative	vajzë	rrufe	dele
	Accusative	vajzë	rrufe	dele
	Dative	vajz-e	rrufej-e	delej-e
	Ablative	vajz-e	rrufej-e	delej-e
	Genitive	i vajz-e	i rrufej-e	i delej-e
Definite	Nominative	vajz-a	rrufej-a	delj-a
	Accusative	vajzë-n	rrufe-në	dele-n
	Dative	vajzë-s	rrufe-së	dele-s
	Ablative	vajzë-s	rrufe-së	dele-s
	Genitive	i vajzë-s	i rrufe-së	i dele-s

Table 2.2

The feminine declension takes the suffix *-e* for indefinite dative, ablative, and genitive. Since feminine nouns are often characterized as ending in a vowel, this leads to some phonological interests. If the stem-final vowel is ë (such as in vajzë), the ë drops, as is expected in most cases, like *vajze* instead of **vajzëe*. According to Newmark et al. (1982), if the vowel is anything but ë, then a *j* is inserted between the stem and suffix, like *deleje*. In the definite, there

is even more interesting phonology associated with stress. The definite nominative suffix, *-a*, causes deletion of *unstressed* vowels, whether it is *ë* or not, but vowels that are not *ë* still take an inserted *j* prior to deletion, like *delja* (Newmark et al. 1982: 161). Stressed vowels do not delete, as before, and also take the inserted *j*. In the accusative (*-në)*, and the dative, ablative, and genitive (*-së*), the suffixes end in *ë*, which deletes when following an unstressed syllable (*vájzë-s*), but will remain when following a stressed syllable (*rrufé-së*).

2.1.2 The Plural Paradigms

In the singular, there are three classes of nouns, corresponding to morphosyntactic gender. It would seem logical, then, that the plural paradigms would follow suit. This is not the case. Take Table 2.3, showcasing the Class I (*masculine)* noun LIS 'oak' and the singular Class III (*feminine*) noun VAJZË 'girl'.

			lis 'oak'	vajzë 'girl'
Singular	Indefinite	Nominative	lis	vajzë
		Accusative	lis	vajzë
		Dative	lis-i	vajz-e
		Ablative	lis-i	vajz-e
		Genitive	i lis-i	i vajz-e
	Definite	Nominative	lis-i	vajz-a
		Accusative	lis-in	vajz-ën
		Dative	lis-it	vajz-ës
		Ablative	lis-it	vajz-ës
		Genitive	i lis-it	i vajz-ës
Plural	Indefinite	Nominative	lis-a	vajz-a
		Accusative	lis-a	vajz-a
		Dative	lis-a-ve	vajz-a-ve
		Ablative	lis-a-sh	vajz-a-sh
		Genitive	i lis-a-ve	i vajz-a-ve
	Definite	Nominative	lis-a-t	vajz-a-t
		Accusative	lis-a-t	vajz-a-t
		Dative	lis-a-ve	vajz-a-ve
		Ablative	lis-a-ve	vajz-a-ve
		Genitive	i lis-a-ve	i vajz-a-ve

Table 2.3

The Class I noun, LIS ‘oak,’ differs from the Class III noun, VAJZË 'girl,' in the singular, but they share an identical plural paradigm. In contrast, the Class I noun MIK ‘friend’ is in a different plural class than LIS, but the same plural class as the Class III noun RRUFE ‘thunderbolt’ (shown in Table 2.4). This indicates that the plural paradigms are independent from the singular paradigms. Table 2.3 showcases one of the plural declensions, which I label Class A. Table 2.4 shows plural Class B.

		rrufe ‘thunderbolts’	miq ‘friends’	peshq ‘fish’
Indefinite	Nominative	rrufe	miq	peshq
	Accusative	rrufe	miq	peshq
	Dative	rrufe-ve	miq-ve	peshq-ve
	Ablative	rrufe-sh	miq-sh	peshqi-sh
	Genitive	i rrufe-ve	i miq-ve	i peshq-ve
Definite	Nominative	rrufe-të	miq-të	peshqi-t
	Accusative	rrufe-të	miq-të	peshqi-t
	Dative	rrufe-ve	miq-ve	peshq-ve
	Ablative	rrufe-ve	miq-ve	peshq-ve
	Genitive	i rrufe-ve	i miq-ve	i peshq-ve

Table 2.4

Class B contains both feminine nouns, such as RRUFE ‘thunderbolt’, and masculine nouns, such as MIK ‘friend’. Masculine nouns in this class that end in *k* palatalize in the plural, hence *miq* ‘friends’ and *mik* ‘friend’. Stems that end in consonant clusters, such as PESHQ (or *ëz* such as NJERËZ ‘people’, *s* such as NXËNËS ‘pupils’, or *ër* such as PRINDËR), will insert (unstressed) *i* before the indefinite ablative and definite nominative and accusative endings (Newmark et al. 1982). As before, the definite nominative and accusative endings do not drop *ë* when immediately following a stressed syllable, so *míqtë* takes the same ending as *péshqit* (underlyingly *-të*). What is most notable is that both plural classes have nearly identical endings except for an additional exponent in Class A (*-a*).

Also notable about these paradigms is indefinite ablative. For almost all nominal paradigms, ablative is syncretic with dative and genitive. In this single instance, ablative is *not* syncretic, taking the form *-sh* rather than the dative and genitive *-ve*. However, as I will discuss more in the next section, adjectives split from nouns in this area, instead remaining syncretic with dative and genitive.

2.1.3 Adjectives

The adjective paradigms are similar to noun paradigms, except for a few differences. Nouns have inherent gender, but adjectives do not. As such, adjectives will often have different stems corresponding to gender, either masculine or feminine. Neuter has fallen almost entirely out of favor[3], replaced by masculine forms (Newmark et al. 1982: 185). Perhaps the most conspicuous of the differing-stem phenomenon is ZI 'black', shown in table 2.5.

			Masculine	Feminine
			zi 'black'	zezë 'black'
Singular	Indefinite	Nominative	i zi	e zezë
		Accusative	të zi	të zezë
		Dative	të zi-u	të zez-e
		Ablative	të zi-u	të zez-e
		Genitive	i të zi-u	i të zez-e
	Definite	Nominative	i zi-u	e zez-a
		Accusative	të zi-un	të zez-në
		Dative	të zi-ut	së zez-së
		Ablative	të zi-ut	së zez-së
		Genitive	i të zi-ut	i së zez-së

Table 2.5

The masculine-agreeing lexeme, ZI, takes the Class I suffix *-i*, but because it ends in a stressed vowel (*zí*), the suffix changes to *-u*. What is also notable is that masculine-agreeing

[3] This is the case for the adjectives themselves, but the adjective concord particles still mark neuter.

adjectives take the (singular) Class I paradigm. The feminine-agreeing adjectives likewise take Class III suffixes.

Adjectives also differ from nouns in their particle of concord, a proclitic that may appear before the adjective (such as *i zi*) that marks case, definiteness, number and gender (Newmark et al. 1982). The particle of concord is quite fascinating and presents an issue to the idea of *second position*. When an adjective precedes the noun in the nominal phrase, the particle of concord falls in first position. However, despite being in first position, the clitic properties do not fall immediately after the particle of concord. It is for this reason that I define second position as after the first *lexical* word.

Another difference is the plural paradigms, which behave similarly to the noun paradigms. Where the singular paradigms are basically identical with their noun counterparts, the plural paradigms are slightly different.

			Masculine	Feminine
			zinj 'black'	zeza 'black'
Plural	Indefinite	Nominative	të zinj	të zez-a
		Accusative	të zinj	të zez-a
		Dative	të zinj-ve	të zez-a-ve
		Ablative	të zinj-ve	të zez-a-ve
		Genitive	i të zinj-ve	i të zez-a-ve
	Definite	Nominative	të zinj-të	të zez-a-t
		Accusative	të zinj-të	të zez-a-t
		Dative	të zinj-ve	të zez-a-ve
		Ablative	të zinj-ve	të zez-a-ve
		Genitive	i të zinj-ve	i të zez-a-ve

Table 2.6

Table 2.6 showcases an interesting phonological allomorphy in the masculine plural. For both adjectives and nouns, if the stem ends in a stressed *i*, *nj* inserts onto the stem (Newmark et al. 1982: 188).

The plural noun paradigms do not take gender into account, with both classes containing a mix of masculine, neuter, and feminine nouns. However, in the adjective paradigms, which are *nearly* identical, they are clearly defined for each gender; Feminine-agreeing adjectives take the Class A paradigm, and masculine-agreeing adjectives take Class B. There is however one noticeable difference between adjectives and nouns here; in the plural indefinite ablative, adjectives remain syncretic with dative and genitive with the form *-ve*, whereas nouns have a different form *-sh*. This is striking because it is the *only* suffix that is different for adjectives than with nouns.

2.1.4 The Paradigms Without Case and Definiteness

In the previous three sections, I discussed the inflectional suffixes that include the properties of case, definiteness, number, and gender. Those paradigms are considerably larger than the paradigms for the wordforms of the subsequent lexical words in the phrase, which mark simply number and gender. The forms devoid of case and definiteness are identical to the indefinite nominative forms, shown in Table 2.7.

	‘lis’ oak	‘rrufe’ thunderbolt	‘zi’ black	‘zezë’ black
singular	lis	rrufe	zi	zezë
indefinite nominative singular	lis	rrufe	zi	zezë
plural	lis-a	rrufe	zinj	zez-a
indefinite nominative plural	lis-a	rrufe	zinj	zez-a

Table 2.7

Nouns in Class A contain the additional exponent of *-a*. As is also expected, nouns in Class B without case and definiteness continue to lack an additional exponent like the one found in Class A.

The clitic’s cumulative expression presents an interesting conundrum for a model of morphology because, as I showed in these sections, lexemes have two different paradigms

depending on their location within the phrase. In Chapter 4, I develop an analysis accounting for these two seemingly separate paradigms.

2.2 The Nominal Phrase

In section 2.2, I detailed the relevant morphological facts regarding the noun phrase. In this section, I detail the empirical facts regarding the placement of the clitic properties within the nominal phrase.

The examples that I have introduced to this point showcasing the distributional facts of the clitic properties, such as (1.1) *vajzat e mira*, have been no larger than a noun and an adjective. As is to be expected, the clitic's distributional facts hold up on noun phrases with more lexical words, such as (2.2).

(2.2) a. vajz-ave të mir-a e të urt-a
girl-DAT.DEF.F.PL DAT.DEF.F.PL good-F.PL and F.PL quiet-F.PL
'good and quiet girls'

b. të mir-ave vajz-a të urt-a
DAT.DEF.F.PL good-F.PL girl-DAT.DEF.F.PL F.PL quiet-F.PL
'good and quiet girls'

In (2.2a), the first lexical word is the noun, *vajzave*, which marks all four properties. The other two lexical words, the adjectives *mira* and *urta*, only mark number and gender. This is the case no matter how many lexical words there are in the phrase.

However, a question that may be raised here is how noun phrases work, such as copula and genitive constructions. When the adjective is split from the noun phrase by the copula, case and definiteness remains unmarked on the predicative adjective (2.3).

(2.3) rrufe-t janë të urt-a
thunderbolt-NOM.DEF.F.PL be.3PL F.PL[4] quiet-F.PL
'the thunderbolts are quiet'

Although genitive noun phrases are set within the noun phrase they modify, the clitic distributions of the two phrases remain separate (2.4).

(2.4) të mëdhenj lis-a i burr-it të
NOM.DEF.M.PL big.NOM.DEF. M.PL oak-M.PL GEN man-GEN.DEF.M.SG GEN.DEF.M.SG
vogël
small.M.SG
'the big oaks of the small man'

As seen in (2.3), genitive phrases are separated from the phrases they modify by a pre-phrasal clitic *i*. The first lexical word immediately following this genitive clitic marks all four properties, but the remaining lexical words only mark number and gender.

2.3 Summary

In this chapter, I detailed the empirical facts surrounding the nominal morphology and the placement of the clitic properties. In Chapter 3, I discuss and develop an analysis of the placement of the clitic properties within the phrase, specifically in second position, that treats the Albanian clitic as edge inflection and maintains its cumulative exponence. In Chapter 4, I discuss how morphology accounts for the cumulative expression and the fact that lexemes have apparently distinct paradigms based on their placement within the phrase.

[4] This is another example of the different distributional facts of the adjective's concord particle. As I discussed in the previous section, the first concord particle in the noun phrase marks all four properties. In examples where the adjective is within the NP, like (1.1) *vajzat e mira*, it is straightforward that the particle marks all four properties, but in (2.3), the actual properties expressed are less obvious because the form is identical to the underlying form *të* rather than *e* like in (1.1), despite being the first concord particle. These distributional facts are certainly worth studying but are beyond the scope of this thesis.

Chapter 3: Placement

In this chapter, I discuss how the cumulative exponence of the Albanian data presents difficulties to the syntactic placement of case and definiteness. The crux of the issue presented to many previous models of 2P clitic placement is in a particular assumption; many of these models assume form is inserted prior to the process that finalizes the clitic in second position. Some, such as Prosodic Inversion (Halpern 1995), go so far as to assume that form licenses placement. However, the Albanian data presented here challenge this assumption due to its cumulative exponence; if form is inserted prior to placement, it is not immediately clear how the two sets of properties become expressed cumulatively.

The assumption that form precedes placement is made in part due to a more underlying assumption; models such as Halpern (1995) and Anderson (2005), assume clitics are self-contained units that do not morphophonologically interact with their host. These kinds of clitics are categorized as *phrasal affixation* (Spencer and Luís 2012), and are more typical than clitics that behave like the Albanian clitic, known as *edge inflection*. Any analysis that assumes clitics are generally phrasal affixes will struggle to account for the Albanian clitic. Thus, to account for the Albanian clitic, an analysis must account for edge inflection in particular.

In 3.1 and 3.2, I detail two models that assume clitics are generally phrasal affixes and demonstrate why they fail when applied to the Albanian clitic. Section 3.1 details Prosodic Inversion (Halpern 1995) and section 3.2 details Anderson (2005)'s Optimality Theory account. I use these approaches because, although they are vastly different, both have the same point of failure. In 3.3, I provide a possible analysis using Head-Driven Phrase Structure Grammar (Pollard and Sag 1994).

3.1 Prosodic Inversion

In this section, I detail a popular analysis using Prosodic Inversion (PI, Halpern 1995). PI is an analysis of 2P clitic placement that assumes the form of the clitic licenses movement to its final position. PI works well with transformational grammars, such as Distributed Morphology (DM; Siddiqi 2019, Embick 2015), because it posits movement.

The basic premise of PI is this; the clitic is placed to the left of the phrase and, after phonological information is inserted, its prosodic deficiency moves it behind a host. What exactly qualifies as a host is not entirely clear, but for the purposes here I assume the first stressed lexical item, or PWord, qualifies. To demonstrate how this works, I use the DM framework.

The crucial assumptions of DM for this analysis are: the syntactic atoms are abstract morphemes rather than words, X-bar syntactic theory, and a realizational model of morphology. The model of DM uses Spell-Out, where the incomplete syntactic derivation splits off from the final Logical Form and goes to Phonological Form. It is here that, according to Harley and Noyer (1999), morphological operations happen, followed by the insertion of phonological information, called Vocabulary Insertion (VI). After VI, phonological operations, such as PI, occur. The X-bar theory model that I use here is the model described by Carnie (2013).

To make this concrete, I use the Serbian 2P clitic that I introduced in Chapter 1, shown in (3.1).

(3.1) čovek =je voleo mariju
man.NOM.SG AUX.3.SG love.PST.M Maria.ACC
'The man loved Maria'

Before the clitic gets to second position, PI places it syntactically to the left of the phrase. How to represent the underlying syntactic position of the clitic to the left of the phrase is not

clear. This is a fact that Halpern (1995) intentionally glosses over, preferring instead to emphasize the final phonologically-motivated position of the clitic in second position. This is not an issue for phrasal affixes, such as the Serbian auxiliary, but becomes an issue for edge inflection examples, such as the Albanian properties. Thus, a simplified tree is shown in Figure 3.1.

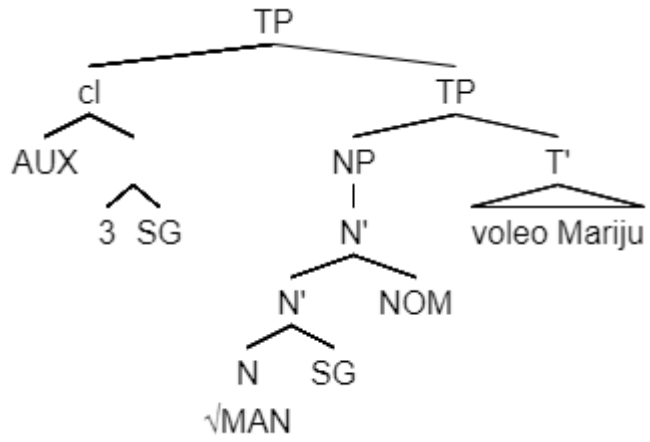

Figure 3.1

In this tree, I purposefully gloss over the part of the phrase to the right of the NP, which has little to do with the clitic, in order to focus on the clitic and its PWord host. Also notable is one of the crucial assumptions of DM; all grammatical properties are their own abstract morpheme. In Figure 3.1, the properties of singular and nominative are separate nodes from each other, but the realized form requires these properties to be packaged together. To do this, DM uses a step during morphological operations called Fusion (Siddiqi 2019). Fusion takes neighboring functional, or grammatical, nodes and fuses them together to create larger complex nodes. This step generates the structure in Figure 3.2.

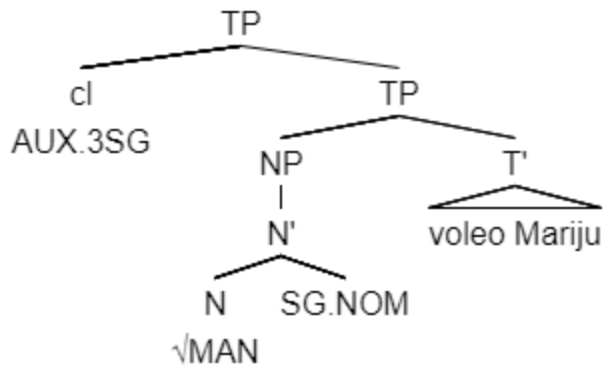

Figure 3.2

Now that morphological operations have ceased and we have fully packaged morphemes, we are ready for VI to occur. Here, phonological material is inserted. The clitic's prosodic deficiency then motivates the movement of the clitic behind the first PWord of the phrase, in this case *čovek*. This process is shown in Figure 3.3.

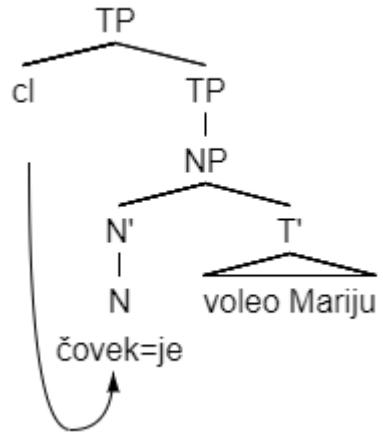

Figure 3.3

While the final result in Figure 3.3 is correct, this approach has some issues. It is not entirely clear what qualifies as the host, nor by what process the clitic is prevented from attaching at a later point in the phrase.

More importantly, these steps do not account for the cumulative expression seen in the Albanian data. While this is not a problem for the Serbian clitic, which is not cumulatively

expressed with the properties expressed by the PWord *čovek*, it is not clear how the Albanian cumulative expression would be handled by PI.

Notable in Figure 3.3 is that the auxiliary particle is not syntactically represented behind *čovek*, but that it is phonologically surfacing in this position. This fact is not entirely relevant in the Serbian example, but it will become more relevant when applied to Albanian. The Fusion step posited by DM merges functional nodes that are syntactic neighbors, but the phonological position of the clitic and its underlying syntactic location are separate. The syntactic position of the Albanian clitic will be too far to fuse with the head containing number and gender, and so the cumulative expression of the Albanian clitic cannot be accounted for. To make this concrete, I will apply PI to the Albanian example.

The clitic properties are initially placed to the left of the phrase, shown in Figure 3.4. Again, I gloss over the remaining part of the phrase that is not relevant to the clitic.

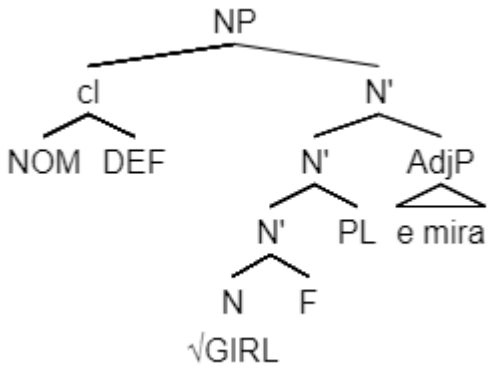

Figure 3.4

Then the Fusion step fuses the neighboring functional nodes into complex nodes. Fusion separately merges the clitic properties together and the affixal properties after the root √GIRL together. This process generates the structure in Figure 3.5.

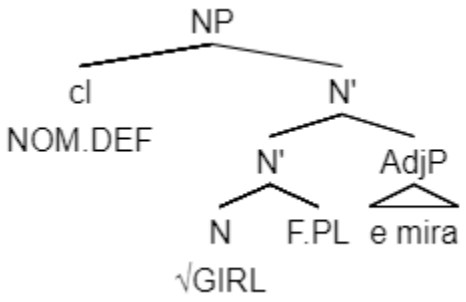

Figure 3.5

At this point, the Fusion step cannot merge the clitic properties and the affixal properties because they are not neighboring nodes. The next step, VI, and the subsequent inversion is shown in Figure 3.6.

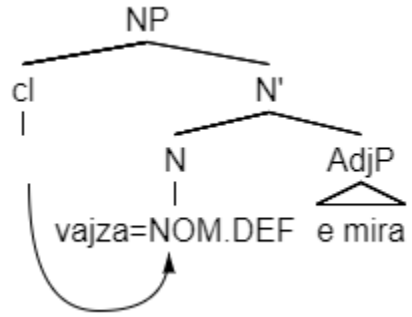

Figure 3.6

Phonologically, the clitic properties neighbor the affixal properties, but their syntactic position remains unchanged. Since the inversion itself is a phonological operation that does not change the underlying syntactic position, Fusion cannot occur after PI either. As shown in Figure 3.6, PI leaves us with two suffixes, one with number and gender and the other with nominative and definite. This runs contrary to the expected outcome of one suffix cumulatively expressing all four properties.

Up until this point I have assumed, as Halpern (1995) lays out, that the final position of the 2P clitic is entirely phonological and is separate from the underlying syntactic position. This

assumption generates an incorrect result, and so perhaps it should be reconsidered. If we abandon this assumption and allow the final position to be both syntactic and phonological, we still run into an issue.

An important part of the VI step is that the actual insertion of phonological material replaces the existing interpretable features (Siddiqi 2019). This means that the morphosyntactic properties contained within the morphemes are entirely replaced by the phonological material. Fusion merges neighboring grammatical nodes, but after VI the nodes no longer retain their interpretable morphosyntactic features. We are thus left with a conundrum; the actual inversion step gets the clitic to neighbor the affixal properties, but the VI step removes the possibility of an additional Fusion step.

As I have demonstrated here, Prosodic Inversion in its current state is ill-equipped to handle 2P clitics that do not behave as phrasal affixes. The clitic and the morphological properties are simply not close enough for a Fusion step to occur, and so we are left with two suffixes. This could possibly be remedied by a raising step before or during VI, such that the clitic is lowered to neighbor the number and gender node. While a solution such as this could in fact allow a Fusion step, the solution would run into issues of motivation. What exactly motivates this lowering step? The answer is not immediately clear, and any answer would make the analysis language-specific in each case. A model that is generalizable such that it can be applied in each case is exceedingly more preferable.

Anderson (2005) argues that the problem with Prosodic Inversion and other rule-based approaches is their rule-based nature. Instead, he proposes an alternative approach using a constraint-based approach, namely Optimality Theory. In the next section, I detail his approach and demonstrate that his approach suffers from a similar problem, leaving us with two suffixes

rather than one. The running problem in both approaches is the assumption that clitics are phrasal affixes, when the Albanian clitic must be treated as edge inflection.

3.2 Anderson (2005) Optimality-Theoretic Approach

The major difference between phrasal affixation and edge inflection is that a clitic categorized as edge inflection morphophonologically interacts with its host. Recall the English example I introduced in Chapter 1, shown here in (3.2).

(3.2) a. The girls' /gɚːlz/ names
b. *The girls's /gɚːlzəz/ names
c. the cheese's /t͡ʃizəz/ aroma

The haplology present in (3.2a) is not present in (3.3c), so we can only assume this haplology is interaction with the stem, like edge inflection. Since the Albanian data showcase rampant allomorphy in the form of large paradigms, the Albanian clitic must be categorized as edge inflection rather than phrasal affixation. To make this concrete, I demonstrate an analysis that assumes clitics are generally phrasal affixes, namely Anderson (2005)'s OT account.

Anderson (2005) argues that cases like the English possessive exponent are actually "interesting evidence in favor of" phrasal affixation (p. 91)[5]. The haplology seen in (3.2a) that is notably absent in (3.2c), Anderson argues, is because "the suffixal /z/ is structurally distinct from stem-final /z/," (p. 93) which allows for easy formulation of reduction from two identical syllabic affixes to a single one. The example he uses is the structure of *dog-s*, shown in Figure 3.7.

[5] It should be noted that Anderson walks the argument for a phrasal-affix only analysis back in Anderson et al. (2006), arguing instead that the English genitive exponent is edge inflection. The analysis sketched in Anderson et al. (2006) is an intriguing one because it seemingly overpredicts the existence of 2P edge inflection. However, as I show in this thesis, the Albanian clitic is a fulfillment of this prediction.

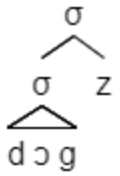

Figure 3.7

The structure of the suffix is not a daughter of the syllable that it attaches to, but a separate syllable itself. Extending this, then, the possessive exponent is an additional syllable of identical phonological material to the plural suffix. This additional syllable is self-contained and does not interact with the stem and present any form of allomorphy and the haplology of the type seen in (3.2a) is simply a regular phonetic reduction. It is this view, that the clitic exponent is self-contained, that causes the analysis to struggle to account for the Albanian clitic.

In the analysis of 2P clitic placement, Anderson (2005) argues against rule-based approaches such as Prosodic Inversion, instead arguing for constraint-based approaches. He develops an analysis of 2P clitics using OT, which is a model that uses constraints to determine the well-formedness of a phrase. Each constraint in OT is violable, so they are organized hierarchically in order to determine which violation is preferred over another. Violations are preferred to occur to constraints lower in the hierarchy, and possible arrangements with the fewest violations (especially of higher ranked constraints) are selected.

Anderson argues that there are two relevant constraints for 2P clitic placement: **NonInitial** and **LeftMost**. The constraint **NonInitial** states that the exponent cannot be at the beginning of the phrase, and **LeftMost** states that the exponent must be as far to the beginning of the phrase as possible. These obviously conflict, and so Anderson ranks **NonInitial** higher than **LeftMost**.

Because Anderson assumes clitics are self-contained, additive phrasal morphology, these constraints work for 2P phrasal affixes, such as the Serbian example, but they do not work for edge inflection like the Albanian case and definiteness properties. Applying this approach to the Serbian example, we get this table:

	NonInitial [AUX]	**LeftMost [AUX]**
je čovek voleo Mariju	*	
☛ čovek je voleo Mariju		*
čovek voleo je Mariju		**
čovek voleo Mariju je		***

Table 3.1

The first construction has a violation in **NonInitial**, so we reject it. The third and fourth constructions have a greater number of violations in the lower ranked **LeftMost** than the second construction, so we reject them as well. As such, we select the second construction, resulting in the expected outcome.

This certainly works for clitics that behave like phrasal affixes, such as the Serbian auxiliary. Since the Serbian auxiliary is a self-contained unit that does not morphophonologically interact with its host, we can categorize it as a phrasal affix. Anderson (2005)'s account assumes clitics are phrasal affixes, so it is no wonder that the Serbian example is adequately explained. However, since Anderson's account generally assumes phrasal affixes, it begins to breakdown when applied to clitics in the edge inflection category.

Applying this approach to the Albanian data, we get this table:

	NonInitial [CASE.DEF]	**LeftMost [CASE.DEF]**
NOM.DEF vajza e mira	*	
☛vajza NOM.DEF e mira		*
vajza e NOM.DEF mira		**
vajza e mira NOM.DEF		***

Table 3.2

Since Anderson (2005) argues that clitics are self-contained, additive phrasal morphology, the case and definiteness exponent must be kept separate in this analysis. Keeping the properties of case and definiteness separate is clearly an issue, as it is not an additional suffix, rather cumulatively expressed with number and gender in *one* suffix. In the table, the second option is selected for the same reasons as the Serbian example, which gets the properties to second position. However, we are left with two suffixes, which is an incorrect result. This is a problem for all analyses that assume clitics are self-contained like the phrasal affix category. Instead, an analysis must assume that clitics may morphophonologically interact with their hosts, like clitics in the edge inflection category.

The cumulative expression of the data is a problem for any model that assumes that clitics are phrasal affixes. In both cases I discussed, we are left with two suffixes for the two property sets. Instead, a model should be equipped to deal with edge inflection. In the next section, I detail a possible analysis using such a model, namely Head-Driven Phrase Structure Grammar (HPSG, Pollard and Sag 1994).

3.3 An Alternative Model

In the previous two sections, I demonstrated the issue with the assumption that clitics are generally phrasal affixes. Models like Prosodic Inversion and Anderson's OT approach fail because they leave us with two suffixes when we simply need one. In order for a model to account for data like the Albanian clitic, it must be equipped to account for clitics that are of the edge inflection category. One such model, which I outline here, is HPSG.

HPSG is a model of syntax whose key feature is that it is lexicalist (Montermini 2019). A lexicalist model, as opposed to a morpheme-based model such as DM, assumes that the primitives of syntax are words. These words are *signs*, shown in Figure 3.8.

$$\begin{bmatrix} \text{PHON } \langle vajzat \rangle \\ \text{SYNSEM} \begin{bmatrix} \text{CATEGORY} & [\text{HEAD } noun] \\ \text{CONTENT} & [\text{INDEX } [\text{Gender: F}, \text{Number: PL}]] \end{bmatrix} \\ \text{DOM} \quad [\text{EDGE } [\text{CATEGORY } [\text{Case: NOM}, \text{Definiteness: DEF}]]] \end{bmatrix}$$

Figure 3.8

Signs have three major features: phonology (PHON), morphosyntactic and semantic content (SYNSEM), and linearization domain features (DOM). HPSG is also a linearization-based grammar, where word order is determined through linearization precedence constraints. There are two levels of linearization constraints, individual to a language (or *rule-level*) constraints, and domain-level constraints (Daniels and Meurers 2004). Domain-level constraints are represented by the domain features in the sign. In dealing with edge inflection clitics, the relevant domain feature list is the unary [EDGE] feature.

At a basic level, HPSG assigns the [EDGE] feature to a sign. All signs with the [EDGE] feature have the clitic properties apportioned to them. However, there is only ever *one* sign that marks the edge inflection properties, and so a precedence constraint must be posited. The particular constraint required is somewhat dependent upon the language; some languages, such as Albanian, apportion the properties to the left edge, and others, such as English, apportion the properties to the right edge. For left edge clitics, that constraint is [EDGE] > X, allowing for one and only one sign to be on the edge, and that it gets assigned to the edge before all other signs. In the reverse, the constraint is X > [EDGE], such that the rest of the phrase is linearized before the sign bearing the [EDGE] feature. It is at this point that the sign has its morphosyntactic property set fully packaged and is ready to be relayed to morphology for exponement.

To make this digestible, let's use the English genitive example. The genitive exponent is also present on the right edge of the phrase, and so it falls within the category of edge inflection.

Since the clitic is located on the right edge of the phrase, the constraint X > [EDGE] is applicable. As such, HPSG assigns a sign the [EDGE] feature, in this case the sign representing *girls*, and apportions the [EDGE] list property set {Case: Genitive}. This sign is shown in Figure 3.9.

$$\begin{bmatrix} \text{PHON } \langle girls' \rangle & & \\ \text{SYNSEM} & \begin{bmatrix} \text{CATEGORY} & [\text{HEAD } noun] \\ \text{CONTENT} & [\text{INDEX } [\text{Number: PL}]] \end{bmatrix} \\ \text{DOM} & [\text{EDGE } [\text{CATEGORY } [\text{Case: GEN}]]] \end{bmatrix}$$

Figure 3.9

Only one [EDGE] feature is assigned per phrase, and so the determiner does not get the [EDGE] feature and the corresponding list applied. This gives us the final phrase structure result in Figure 3.10.

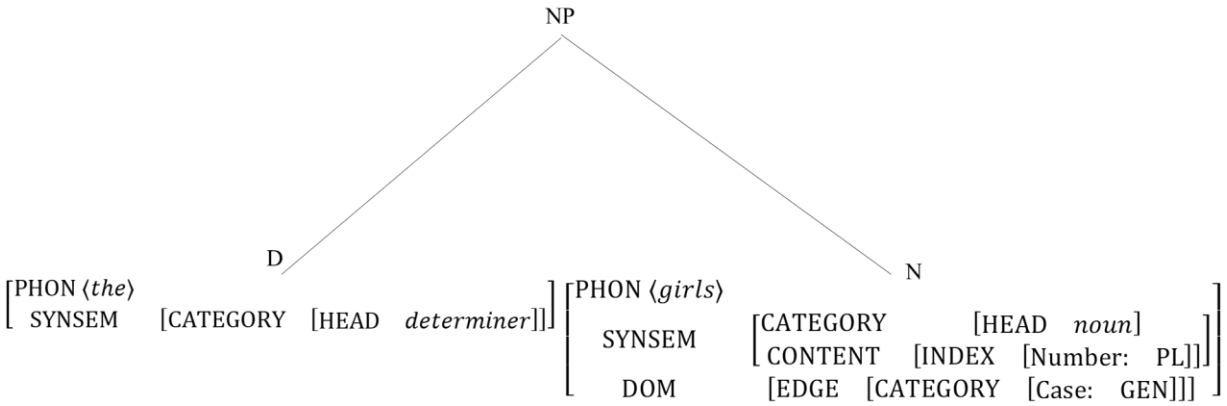

Figure 3.10

The Albanian clitic works in much the same way. Where English also displays cumulatively-expressed edge inflection, the clitic is on the *right* edge of the phrase. The Albanian clitic is on the *left* edge, shown in (3.4)

(3.4) burr-i i zi
man-NOM.DEF.SG.M NOM.DEF.SG.M wretched.M.SG
'the wretched man'

The process for Albanian is much the same as the process detailed above for English, but there is one change; the precedence constraint is [EDGE] > X. This assigns one sign the [EDGE]

feature and linearizes it on the *left* side of the phrase. The EDGE feature list then apportions the EDGE properties. Applying this constraint to Albanian results in the sign shown in Figure 3.11.

$$\begin{bmatrix} \text{PHON } \langle burri \rangle \\ \text{SYNSEM} \begin{bmatrix} \text{CATEGORY} & [\text{HEAD } noun] \\ \text{CONTENT} & \left[\text{INDEX } \begin{bmatrix} \text{Gender:} & \text{M} \\ \text{Number:} & \text{SG} \end{bmatrix}\right] \end{bmatrix} \\ \text{DOM} \quad \left[\text{EDGE } \left[\text{CATEGORY } \begin{bmatrix} \text{Case:} & \text{NOM} \\ \text{Definiteness:} & \text{DEF} \end{bmatrix}\right]\right] \end{bmatrix}$$

Figure 3.11

Another bit of data that separates the English data and the Albanian data is that the Albanian clitics are in second position, whereas the English data are entirely on the right edge. While this initially seems to present an issue, this is easily solved. As argued by most clitic models (Spencer and Luís 2012), 2P clitics are underlyingly proclitics, which means they occur before the phrase. This proclitic nature is readily seen in Halpern (1995)'s Prosodic Inversion[6], which argues that the clitics are placed to the left of the phrase by syntactically trivial means, and it is only prosodic requirements that move the clitic to second position. Edge inflection is no different in this regard; the Albanian data are underlyingly proclitic, as they are placed to the left edge of the phrase. Once it is placed on the left edge, it is then morphology's job to realize the clitic material as an affix. This means that, should morphology realize the clitic as a prefix, it would retain its straightforward proclitic nature. However, since it is realized as a *suffix*, this places it after the first word, or in second position. The same is true about English's genitive exponent. Morphology realizes it as a suffix, which generates as a straightforward enclitic, but if it had realized as a prefix, its position would be the inverse of the Albanian data, or penultimate position.

[6] This is also explicit in Anderson et al. (2006), which seemingly overpredicts that 2P edge inflection such as the Albanian data should exist because 2P edge inflection had yet to be attested.

The analysis presented here ultimately delineates the roles of syntax and morphology into placement and form respectively; syntax locates the edge of the phrase and applies the corresponding morphosyntactic property set, while the form is realized by morphology giving the clitic its final position relative to the word at the edge of the phrase.

3.4 Summary

In this chapter, I introduced the issue presented to many models of clitic placement due to the data's cumulative exponence. Many models assume the form of the clitic is exponed prior to its placement into 2P, which generates two separate suffixes rather than one cumulatively-expressed suffix. The assumption that exponement happens prior to placement is due to a general assumption that 2P clitics are phrasal affixes. In order to account for the Albanian clitic, I show that a model must be able to treat it as edge inflection. The example model I used to analyze the data is Head-Driven Phrase Structure Grammar (Pollard and Sag 1994).

In the next chapter I introduce how an inferential-realizational model, such as Paradigm Function Morphology (PFM, Stump 2001), accounts for the realization of form now that the morphosyntactic property sets have been properly packaged together.

Chapter 4: Insertion of Form

In the previous chapter, I discussed the role of syntax with regard to the Albanian data. Where that was a role of placement, the role of morphology is one of exponement. In this chapter, I demonstrate how morphology deals with the cumulative exponence of the Albanian clitic. Where cumulative exponence posed an issue to models of 2P clitic placement, it notably does not pose an issue for the exponement of the clitic's form.

The cumulative expression of the clitic creates an interesting situation; the wordforms of the first lexical word in the phrase takes a much different, larger paradigm than the wordforms of the subsequent lexical words. Take the paradigms for MIK 'friend.'

			mik 'friend'
Singular	Indefinite	Nominative	mik
		Accusative	mik
		Dative	mik-u
		Ablative	mik-u
		Genitive	i mik-u
	Definite	Nominative	mik-u
		Accusative	mik-un
		Dative	mik-ut
		Ablative	mik-ut
		Genitive	i mik-ut
Plural	Indefinite	Nominative	miq
		Accusative	miq
		Dative	miq-ve
		Ablative	miq-sh
		Genitive	miq-ve
	Definite	Nominative	miq-të
		Accusative	miq-të
		Dative	miq-ve
		Ablative	miq-ve
		Genitive	i miq-ve

Table 4.1

	mik 'friend'
Singular	mik
Plural	miq

Table 4.2

Table 4.1 displays the paradigm for the wordforms of MIK 'friend' when in first position, and Table 4.2 displays the paradigms for when it is in subsequent positions. These paradigms appear separate from each other, which is interesting because we do not normally think of a lexeme as having two different paradigms. Initially, it would seem logical to separate the two paradigms into two different realizational sets, but this is not necessarily required.

In 4.1, I provide an argument for why a single realizational set is preferred over multiple sets. In 4.2, I briefly describe Paradigm Function Morphology (PFM, Stump 2001) as a means for analyzing the Albanian clitic, and in 4.3, I provide a full analysis of the data using PFM.

4.1 One Set, or Two?

The cumulative expression of the clitic creates varying amounts of property expression within the phrase, where the first lexical word marks a larger property set than subsequent lexical words. This property expression can be viewed in one of two ways: either morphology realizes two sets of exponents depending on location within the phrase, or it contains one set for both positions. The first proposal posits that morphology is sensitive to the placement of properties within the phrase and is able to determine which set of exponents is required based on location. The second argument assumes morphology is blind to placement and instead realizes exponents based on the property sets relayed by the syntax. Here, I argue that the latter argument is to be preferred.

First, I argue that one set of exponents is possible. Suffixes in the Albanian data take exponents from one of two possible morphosyntactic property sets; one set includes number, gender, case, and definiteness, whereas the second contains only number and gender. Notable

about these two sets is that one is a proper subset of the other; number and gender are included in *both* sets, but case and definiteness is only in the *first*. Thus, the second set fits neatly within the first set. This demonstrates that a one set analysis is at least as possible as one assuming multiple sets.

Second, I argue that an analysis assuming one set is to be preferred. For morphology to have multiple sets of exponents, it would have to be privy to phrasal placement. This blurs the line between syntax, which contends with placement, and morphology, which contends with form. However, if morphology has one set of exponents, the line would be more cleanly delineated; syntax places words and their morphosyntactic information, then relays morphology for realization to be inserted back into the syntax. Stump (2001) provides a number of reasons to prefer Word-and-Paradigm models, or what he typologizes as *inferential-realizational* models, which assume this cleaner delineation. Some of these reasons I enumerate here.

Stump (2001) argues that realizational theories are preferred over incremental theories because morphology is often not additive. Incremental theories posit that words acquire property sets by the addition of morphemes, which is incompatible with extended exponence, in which a wordform has multiple exponents expressing the same property. Stump also argues that inferential theories are preferred over lexical[7] theories because lexical models must make a distinction between concatenative (i.e. affixal) inflection and nonconcatenative inflection, between which there is no theoretically significant difference. As such, an analysis using an inferential-realizational model, and thus the one-set proposal, is to be preferred.

[7] It is worth noting here the difference between Stump (2001)'s use of the word lexical and Montermini (2019)'s usage. I have here differentiated the terms as lexicalist for the syntactic models that posit primitives are words and lexical for the morphological models that posit abstract morphemes are listed in the lexicon separate from their stems.

Assuming one set of exponents for the whole phrase presents an interesting issue. Since affixes containing case and definiteness are simply suffixes with more information, that would mean multiple realizational rules would be applicable. Take (4.1) for example.

(4.1) të mëdhenj-ve lis-a
DAT.DEF.M.PL big-DAT.DEF.M.PL oak-M.PL
'(for) the big oak'

In (4.1), the adjective *mëdhenj-* has a suffix representing dative, definite, masculine, and plural. The noun *lis-* has a suffix representing masculine and plural, which is a subset of the adjective's properties. Because the same properties on the noun are on the adjective, too, what process selects one suffix over another? PFM has a solution.

4.2 Paradigm Function Morphology

Inferential-realizational models, such as PFM, are particularly as a morphological model paired with lexicalist syntactic models such as HPSG. HPSG packages all of the properties into one sign, forming one cumulative property set.

In order to realize the proper form, PFM uses a series of *realizational rules*, which are rules that specify an output of form based on the morphosyntactic property input.

(4.2) RR_1 ⟨X, {Number: PL, Gender: F}⟩ → Xa

The example in (4.2) says that RR_1, or Realizational Rule 1, states that for the input of a stem X paired with the morphosyntactic properties of plural and feminine, the output is the suffix *-a*. Realizational rules are often in competition with one another. To capture this phenomenon, realizational rules are placed in groups called rule blocks, where each rule block contains rules that compete with one another for application. Thus, only one rule from each rule block applies.

Rules are eligible for application if the rule's property set is a subset of the input word's property set. To explain, take (4.3).

(4.3)
RR_1 ⟨X, {Number: PL, Gender: F}⟩ → Xa
RR_2 ⟨X, {Number: PL, Gender: F, Case: NOM, Definiteness: DEF}⟩ → Xat

In this example, there are two realizational rules, both containing plural and feminine. If we test (4.3) with ⟨*vajzë*, {Number: PL, Gender: F, Case: NOM, Definiteness: DEF}⟩, *both* rules are applicable, since the sets of both rules are subsets of the property sets of the input word, *vajzë*. This is the sense in which the rules compete.

When multiple eligible rules are competing for application, PFM selects which rule to apply based on Pāṇinian ordering. Pāṇinian ordering is a specificity-based ordering where the narrowest applicable rule, or the rule with the most specific applicable property set, is selected and applied. In our example above, both rules are applicable. Since RR_2 is more specific than RR_1, RR_2 is selected and applied, giving us the correct form *vajzat*.

Blocks are ordered, as well, according to the paradigm function. In the case of Albanian, that function is displayed in (4.4).

(4.4) $PF(\langle X, \sigma\rangle) = (NaR_2(NaR_1(\langle X, \sigma\rangle)))$

The function shows in which order rule blocks apply their rules. In (4.4), the first block, labeled *NaR_1*, applies first and the second block, *NaR_2*, applies after. In the case that the rules from each block are both suffixes or both prefixes, this would result in the rule from the first block applying closer to the stem than the second.

4.3 Accounting for Syncretism Between the Paradigms

In this section, I build an analysis using PFM. Such an analysis has two major problems to consider. First, the analysis must account for the apparent syncretism between the two paradigms, where the sans-case and definiteness wordforms are identical to their indefinite nominative counterparts. This problem is solved using underspecification, which results in a unification of the two paradigms under one set of rules, allowing morphology to remain blind to syntax.

Second, the analysis must account for the fact that lexemes, when in first position, mark case and definiteness, but do not mark these properties in subsequent positions. This problem is one of specificity, where there are multiple competing applicable rules that have different amounts of specification. This can be solved using specificity-based ordering. However, since the Identity Function Default is used to solve the first problem, this second problem melts away completely. A discussion of the full analysis can be found in the Appendix.

Upon first inspection, it would seem the two paradigms are totally separate. However, upon further investigation it becomes obvious that there is some striking similarity, shown in Table 4.3.

	‘lis’ oak	‘rrufe’ thunderbolt	‘zi’ black	‘zezë’ black
singular	lis	rrufe	zi	zezë
indefinite nominative singular	lis	rrufe	zi	zezë
plural	lis-a	rrufe	zinj	zez-a
indefinite nominative plural	lis-a	rrufe	zinj	zez-a

Table 4.3

As Table 4.3 shows, indefinite nominative wordforms are identical to the wordforms that do not mark case and definiteness. The plural form for LIS ‘oak’ is *lisa*, and its indefinite nominative plural form is also *lisa*. It would seem, then, that we could capture this phenomenon by underspecifying indefinite nominative. Such a rule is shown in (4.5).

(4.5) **RR$_1$** ⟨X, {Number: PL, Class: A}⟩ → Xa

This rule applies to LIS when it contains indefinite, nominative, and plural, and when it contains just plural. As such, we can parsimoniously capture both of these wordforms by the same rule. By positing rules like **RR$_1$**, morphology does not require knowledge of placement; words that are not in first position are accounted for by the same set of rules that account for the words in first position. As I discussed in 4.2, this analysis is preferable because it is preferable for morphology to remain blind to syntactic placement.

However, if we try to extend rules like **RR$_1$** to plural Class B, we come across an interesting development. When the wordforms of plural Class B, such as RRUFE 'thunderbolt', mark indefinite nominative plural, the output is identical to the stem. If we were to posit a rule, it would appear as in (4.6).

(4.6) **RR$_2$** ⟨X, {Number: PL, Class: B}⟩ → X

However, by taking full advantage of underspecification, we can use what is known as the Identity Function Default (IFD). The IFD is where, if no rule from a block applies, the output is identical to the stem. This means we simply forego a rule for plural Class B when the word marks either indefinite nominative plural or simply plural, allowing the IFD to apply instead.

In this analysis, the IFD has a large amount of utility because the indefinite nominative singular wordforms, as well as the simply singular wordforms, are also identical to their stems (see Table 4.3). This means we can forego rules for any of the singular forms without case and definiteness, regardless of class. We can also forego a rule for indefinite nominative singular. In any of these cases, the IFD would apply, outputting a form identical to the stem. Thus, the only rule specified for any of the wordforms in the paradigm for words not in first position is for plural Class A.

Using the IFD to capture plural Class B has interesting implications. Recall that plural Class A and plural Class B are differentiated only by Class A's additional exponent *-a*, and share the same paradigm for the remaining exponents, shown in Table 4.4.

		Plural Class A	Plural Class B
		lis 'oak'	rrufe 'thunderbolt'
Indefinite	Nominative	lis-a	rrufe
	Accusative	lis-a	rrufe
	Dative	lis-a-ve	rrufe-ve
	Ablative	lis-a-sh	rrufe-sh
	Genitive	i lis-a-ve	i rrufe-ve
Definite	Nominative	lis-a-t	rrufe-të
	Accusative	lis-a-t	rrufe-të
	Dative	lis-a-ve	rrufe-ve
	Ablative	lis-a-ve	rrufe-ve
	Genitive	i lis-a-ve	i rrufe-ve

Table 4.4

Beyond the additional exponent *-a*, LIS 'oak' and RRUFE 'thunderbolt' have the same exponent for dative plural, *-ve*, definite accusative plural, *-t*(*ë*), and so on. This allows us to underspecify plural for the plural paradigms beyond indefinite nominative. Essentially, I argue here that the Class A exponent *-a* is not in competition with the other plural exponents. This lack of competition requires these exponents to be in separate rule blocks.

This has direct implications on the issue of the varying levels of property expression. Essentially, we should expect rules within the analysis to have different levels of specificity due to the differing amounts of property expression; sometimes lexemes mark case and definiteness, and other times they do not. However, using the IFD to capture the singular forms and the plural Class B forms that do not also express case and definiteness eliminates a large chunk of this issue. To make this concrete, if we try to apply the analysis to RRUFE 'thunderbolt' while it is marking indefinite dative singular, we should expect two rules to apply, shown in (4.7).

(4.7)

- **RR$_3$** ⟨X, {Number: SG, Class: III}⟩ → X
- **RR$_4$** ⟨X, {Definiteness: NDEF, Case: DAT, Number: SG, Class: III}⟩ → Xe

We would initially expect both of these to appear in the analysis because one rule gives a form for when RRUFE is marking case and definiteness and another for when RRUFE is not. However, **RR$_3$** is actually captured by the IFD, which means it cannot be applicable to RRUFE. Essentially, a lexeme will never have different levels of specificity in its applicable rules due to the different levels of property expression in the phrase.

This is not to say that underspecification does not create issues of different levels of specificity within the blocks; this is undoubtedly present due to the syncretism between singular Class I and singular Class II, where these two classes mark a different exponent from Class III. To make this concrete, see (4.8).

(4.8)

- **RR$_5$** ⟨X, {Definiteness: NDEF, Case: DAT, Number: SG, Class: III}⟩ → Xe
- **RR$_6$** ⟨X, {Definiteness: NDEF, Case: DAT, Number: SG}⟩ → Xi

Applying the rules in (4.8) to our earlier example of RRUFE, which is a Class III noun, we run into a little bit of an issue. Both **RR$_5$** and **RR$_6$** are applicable, since **RR$_5$** is just a more specific version of **RR$_6$**. In such a case, Pāṇinian ordering dictates that the more specific of the applicable rules applies. Thus, **RR$_5$** is selected and applies, giving us *rrufeje*. If we apply these rules to a word in Class I, such as LIS ‘oak’, we get a different result. Since LIS does not contain {Class: III}, **RR$_5$** is not applicable. **RR$_6$**, thus being the only applicable rule, is selected and applied. This leaves us with *lisi*.

While this is useful for when there is varying levels of specificity within the rules themselves, this is notably not applicable for the different levels of property expression due to the employment of the IFD. In the case where a lexeme is marking plural Class A, which, in

contrast to Class B, has its own rule, this is still notably not an issue. Essentially, when a plural Class A word is marking case and definiteness, it takes its first exponent (*-a*) from Block 1 and its second exponent from Block 2. When that same word is not marking case and definiteness, the same rule in Block 1 applies, applying the same exact exponent (*-a*), but none of the rules in Block 2 apply. This means its case and definiteness wordforms are not in direct competition with its wordforms devoid of case and definiteness.

In this section, I developed an analysis of the property expression within the Albanian noun phrase using PFM. This analysis accounts for the syncretism between the paradigms in first position marking case and definiteness and the paradigms in subsequent positions not marking case and definiteness using underspecification and the IFD. A natural result of employing the IFD in such a way also solves the issue of lexemes having varying levels of property expression depending on position in the phrase, ensuring that the rules applying to the lexemes marking case and definiteness are not in competition with the rules applying to the lexemes devoid of case and definiteness. The full analysis can be found in the Appendix.

4.4 Summary

In this chapter, I developed an analysis of the Albanian data, showcasing how cumulative exponence, which poses an issue for many syntactic models of 2P clitic placement, poses no issue for morphology. The model used was Paradigm Function Morphology because of its utility as an inferential-realization model and the direct parallels with lexicalist models of syntax such as Head-Driven Phrase Structure Grammar. Employing underspecification and the Identity Function Default, the issues posed by cumulative exponence through the varying levels of property expression are solved.

Chapter 5: Conclusion

Albanian noun phrases contain a second-position clitic that exhibits cumulative exponence. This cumulative exponence, where a single suffix contains two sets of properties placed by different means, presents an issue for many models of second-position clitic placement. The models for which the Albanian data present a problem, such as Prosodic Inversion and Anderson (2005)'s Optimality Theory account, all generally assume that clitics are self-contained units that do not interact with their host. This is not an unreasonable assumption, as many clitics are of this type, referred to as phrasal affixation. However, not all clitics behave in such a unitary way; the English possessive and the Albanian case and definiteness properties morphophonologically interact with their host, displaying allomorphy similar to affix-stem combinations. This category of clitic is known as edge inflection.

In order to adequately account for the data, a 2P clitic placement model must be able to treat clitics as edge inflection. The model I used as an example in Chapter 3 is Head-Driven Phrase Structure Grammar. In Chapter 3, I presented an argument in favor of lexicalist models like HPSG. Since the data exhibit cumulative exponence, the model must package all morphosyntactic property sets into one single unit before it can be relayed to morphology. To do so, HPSG uses a linearization domain precedence constraint [EDGE], which is assigned to a lexical word, which linearizes that sign to the appropriate edge. The appropriate edge is determined at the level of the language applying the feature to either the left edge—[EDGE] > X—or the right edge—X > [EDGE]; in Albanian, the appropriate edge is the *left* edge. Then, once the left edge has been determined, the clitic properties are apportioned, and a fully packaged sign is ready for exponement.

A problem for further research is properly defining what "second position" truly is. This is a problem that many have grappled with (Halpern 1995), and one I gloss over due to the presence of the adjectival particle of concord and the genitive pre-phrasal particle, which are often positioned before the word marking case and definiteness. In order to get around this issue, I was forced to state that the first *lexical* word of the phrase could bear the case and definiteness clitic.

I also argue for inferential-realizational models of morphology such as PFM because exponement cannot precede placement, which is a base assumption in models like PI and Anderson's OT account. Under a model like PFM, the morphosyntactic property sets license the form, which means the property sets must be fully packaged and available at the time of realization. In my analysis, this is accomplished using HPSG, which then relays that information to morphology for realization.

The cumulative expression of the phrase presents a couple interesting problems to morphology in the varying levels of property expression within the phrase. This issue creates two paradigms for lexemes based on their location, which is not how paradigms are traditionally analyzed. Since it is preferable for morphology to remain blind to syntactic placement, my analysis must make use of one set of rules for *both* sets of paradigms. Interestingly, the paradigms are not as separate as it would initially seem; indefinite nominative singular forms are syncretic with their purely singular counterparts. This is another conundrum that morphology must account for. Both of these problems, the varying levels of property expression and the syncretism between paradigms, are solved using underspecification and the Identity Function Default.

This analysis is quite interesting, since it does not initially seem like it would be extendable to other languages. If a language has a similar situation where there are varying levels of property expression based on syntactic position, but the syncretism between paradigms is not present, it is not entirely clear how morphology would solve this issue using an analysis like mine. However, my analysis could still be used for situations like these since it still employs specificity-based ordering. It just so happened that, in Albanian at least, this turned out to not be necessary.

One area that would be very interesting to continue research is on the Albanian adjectival particle of concord. In my analysis, I focused solely on the lexical words within the phrase. However, the particle of concord may also mark all four properties. Interestingly, though, the particle of concord changes the properties that it expresses based on its position within the phrase with regard to other particles of concord. Essentially, the first particle of concord, whether its adjective precedes or postcedes the noun, marks all four properties. Then subsequent particles of concord have a different set of properties. It would be interesting to see if my analysis may extend to the particles of concord as well, or if additional work may be required.

This thesis contributes to the literature by arguing for a certain type of model. In order to account for the data, morphology and syntax must be more cleanly delineated than what models like Distributed Morphology posit. This specifically argues for lexicalist models of syntax, like Head-Driven Phrase Structure Grammar and inferential-realizational models of morphology, like Paradigm Function Morphology. By cleanly delineating the roles of syntax and morphology, we can specify exactly what those roles are. In this thesis, I argue the role of syntax is in the placement of morphosyntactic properties within the phrase. Morphology's role, then, is in the realization of the form of the morphosyntactic properties. This type of analysis yields an

interesting result for edge inflection clitics. Ultimately, syntax is placing the clitic properties in first position with the first word, arguing that the clitic is underlyingly proclitic. It is then morphology that determines the clitic to be in second position due to its realization as a suffix. Had it instead been realized as a prefix, then it would remain in its underlyingly proclitic position. This can be extended for edge inflection that falls on the right edge, such as the English genitive exponent, arguing that these therefore must be underlyingly enclitic. I know of no penultimate clitics, but the argument I present here is that, if they were found, they would be underlyingly enclitic that were morphophonologically realized as a prefix.

References

Anderson, Stephen. 2005. *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.

Anderson, Stephen, Lea Brown, Alice Gaby, and Jacqueline Lecarme. 2006. Life on the Edge: There's Morphology There After All! *Lingue e Linguaggio* 1: 33-48.

Daniels, Michael and Detmar Meurers. 2004. GIDLP: A Grammar Format for Linearization-based HPSG. In *Proceedings of the HPSG04 Conference Center for Computational Linguistics*, edited by Stefan Müller. Stanford, CA: CSLI Publications.

Embick, David. 2015. *The Morpheme: A Theoretical Introduction.* Boston/Berlin: De Gruyter Mouton.

Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics*. Stanford, CA: CSLI Publications.

Kuhn, Jonas. 2012. Interfaces in Constraint-Based Theories of Grammar. In *the Oxford handbook of linguistics interfaces*, edited by Gillian Ramchand and Charles Reiss, 613-650. Oxford: Oxford University Press.

Montermini, Fabio. 2019. Later generative grammar and beyond: Lexicalism. In *the Oxford handbook of morphological theory*, edited by Jenny Audring and Francesca Masini, 122-142. Oxford: Oxford University Press.

Newmark, Leonard, Philip Hubbard, and Peter Prifti. 1982. *Standard Albanian: A Reference Grammar for Students*. Stanford, CA: Stanford University Press.

Newmark, Leonard. 1998. *Albanian-English Dictionary*. Oxford: Oxford University Press.

Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure*. Chicago, IL: The University of Chicago Press.

Prince, Alan, and Paul Smolensky. *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick: Rutgers University Press, 1993.

Siddiqi, Daniel. 2019. Distributed Morphology. In *the Oxford Handbook of Morphological Theory*, edited by Jenny Audring and Francesca Masini, 143-164. Oxford: Oxford University Press.

Sims, Andrea and Brian Joseph 2018. Morphology versus syntax in the Balkan verbal complex. *Balkan Syntax and (Universal) Principles of Grammar*, 285.

Spencer, Andrew and Ana R. Luís. 2012. *Clitics: An Introduction*. Cambridge: Cambridge University Press.

Spencer, Andrew and Ana R. Luís. 2012. The Canonical Clitic. In *Canonical Morphology and Syntax,* edited by Dunstan Brown, Marina Chumakina, and Greville G. Corbett, 123-150.. Oxford: Oxford University Press.

Stump, Gregory. 2001. *Inflectional Morphology*. Cambridge: Cambridge University Press.

Zwicky, Arnold, and Geoffrey Pullum. 1983. Cliticization vs. inflection: English n't. *Language*, 59(3): 502–513.

Appendix

Block 1

- **RR1.1** ⟨X, {Number: PL, Class: I}⟩ → Xa
- **RR1.2** ⟨X, {Number: SG, Case: DAT}⟩ → Xi
- **RR1.3** When σ = {Number: SG, Case: ABL}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Case: DAT}⟩
- **RR1.4** When σ = {Number: SG, Case: GEN}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Case: DAT}⟩
- **RR1.5** ⟨X, {Number: SG, Class: III, Case: DAT}⟩ → Xe
- **RR1.6** When σ = {Number: SG, Class: III, Case: ABL}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Class: III, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Class: III, Case: DAT}⟩
- **RR1.7** When σ = {Number: SG, Class: III, Case: GEN}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Class: III, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Class: III, Case: DAT}⟩
- **RR1.8** ⟨X, {Number: SG, Definiteness: DEF, Case: DAT}⟩ → Xit
- **RR1.9** When σ = {Number: SG, Definiteness: DEF, Case: ABL}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Definiteness: DEF, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Definiteness: DEF, Case: DAT}⟩
- **RR1.10** When σ = {Number: SG, Definiteness: DEF, Case: GEN}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Definiteness: DEF, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Definiteness: DEF, Case: DAT}⟩
- **RR1.11** ⟨X, {Number: SG, Class: III, Definiteness: DEF, Case: DAT}⟩ → Xi
- **RR1.12** When σ = {Number: SG, Class: III, Definiteness: DEF, Case: ABL}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Class: III, Definiteness: DEF, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Class: III, Definiteness: DEF, Case: DAT}⟩
- **RR1.13** When σ = {Number: SG, Class: III, Definiteness: DEF, Case: GEN}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Class: III, Definiteness: DEF, Case: DAT}⟩) = ⟨Y, σ/{Number: SG, Class: III, Definiteness: DEF, Case: DAT}⟩
- **RR1.14** ⟨X, {Number: SG, Definiteness: DEF, Case: NOM}⟩ → Xi
- **RR1.15** ⟨X, {Number: SG, Definiteness: DEF, Case: ACC}⟩ → Xin
- **RR1.16** ⟨X, {Number: SG, Class: II, Definiteness: DEF, Case: NOM}⟩ → Xt
- **RR1.17** When σ = {Number: SG, Class: II, Definiteness: DEF, Case: ACC}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: SG, Class: II, Definiteness: DEF, Case: NOM}⟩) = ⟨Y, σ/{Number: SG, Class: II, Definiteness: DEF, Case: NOM}⟩
- **RR1.18** ⟨X, {Number: SG, Class: III, Definiteness: DEF, Case: NOM}⟩ → Xa
- **RR1.19** ⟨X, {Number: SG, Class: III, Definiteness: DEF, Case: ACC}⟩ → Xnë

Block 2

- **RR$_{2.1}$** ⟨X, {Number: PL, Case: DAT}⟩ → Xve
- **RR$_{2.2}$** ⟨X, {Category: NOUN, Number: PL, Definiteness: NDEF, Case: ABL}⟩ → Xsh
- **RR$_{2.3}$** ⟨X, {Number: PL, Definiteness: DEF, Case: NOM}⟩ → Xtë
- **RR$_{2.4}$** When σ = {Number: PL, Definiteness: DEF, Case: ACC}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: PL, Definiteness: DEF, Case: NOM}⟩) = ⟨Y, σ/{Number: PL, Definiteness: DEF, Case: NOM}⟩
- **RR$_{2.5}$** When σ = {Number: PL, Case: ABL}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: PL, Case: DAT}⟩) = ⟨Y, σ/{Number: PL, Case: DAT}⟩
- **RR$_{2.6}$** When σ = {Number: PL, Case: GEN}, RR(⟨X, σ⟩) = ⟨Y, σ⟩, where NaR(⟨X, σ/{Number: PL, Case: DAT}⟩) = ⟨Y, σ/{Number: PL, Case: DAT}⟩