

THE MORPHOSYNTAX OF TAGALOG CLITICS: A TYPOLOGICALLY  
DRIVEN APPROACH

A Dissertation

Presented to the Faculty of the Graduate School  
of Cornell University

In Partial Fulfillment of the Requirements of the Degree of  
Doctor of Philosophy

by

Daniel Aaron Kaufman

May 2010

© 2010 Daniel Aaron Kaufman

THE MORPHOSYNTAX OF TAGALOG CLITICS: A TYPOLOGICALLY  
DRIVEN APPROACH

Daniel Aaron Kaufman Ph.D.

Cornell University 2010

In this dissertation, I investigate Tagalog second position clitics in a new empirical depth concluding that these elements are best treated as “non-syntactic” adjuncts to phrase structure. This approach is thus most similar to Anderson’s (1992 *et seq*) “phrasal affix” analysis of clitics but differs in several important respects. Second position clitics are understood here as the spell out of feature bundles which are adjoined to phrase structure in the syntax via an operation Merge (Feature). Unlike bona-fide syntactic heads which are introduced via Merge (Terminal), these feature bundles are not housed under a terminal node within the phrase structure. Nonetheless, limited interactions with syntactic structure are possible. For instance, second position clitics are conjoinable in Tagalog under special conditions and must comply to a locality restriction with the predicate phrase from which they receive their theta-role, two interactions which are highly unexpected on a purely morphological understanding of these elements.

I also review Klavans’ (1980 *et seq*) clitic typology paying special attention to clitic type. I conclude that, once unambiguous morphosyntactic sister clitics (typically syntactic heads) are separated out, two new generalizations emerge: (i) sister clitics can only be manipulated by bona-fide syntactic movement whereas non-sister clitics can be displaced due to phonological factors, (ii) sister clitics can attach phonologically to or away from their complements whereas non-sister clitics cannot attach phonologically to an element outside their syntactic domain. These findings revise our understanding of the repertoire of available clitic positions and attachments,

leading to a welcome simplification. Syntax and phonology are thus not seen to be independent with regard to clitics, as previously claimed (Klavans 1980 *et seq*).

A wide range of syntactic facts concerning Tagalog clitics are also covered here for the first time leading to a new characterization of clitic impenetrable constituents. Impenetrability is shown to be a property of pronominal argument clitics rather than adverbial clitics and requires a locality relation between pronominal arguments and the predicates with which they are associated. This locality condition accounts for the full range of data in a way that cannot be done by previous approaches.

## BIOGRAPHICAL SKETCH

Born in New York City, Daniel Kaufman spent the first two years of his higher education in Hunter College studying Anthropology and Linguistics. There, he developed a strong interest in Austronesian languages which motivated him to study in the University of the Philippines, Diliman, where he ultimately finished his B.A. Having spent three years in the Philippines investigating Tagalog and other Philippine languages, he entered the Ph.D. program at Cornell University with a focus on morphology, typology and Austronesian historical linguistics. During this period, Daniel also had the opportunity to undertake fieldwork on several languages of Sulawesi, Indonesia, investigating clisis and agreement phenomena.

Since the Spring semester of 2009, Daniel has been teaching at the CUNY Graduate School and is currently working to launch the Urban Fieldstation for Linguistic Research, a center for language documentation in the heart of New York City.

Although I may argue against the First Daughter approach to Tagalog clitic placement in §5.9.1, I find it entirely appropriate for the dedication of this thesis:

For Dahlia

## ACKNOWLEDGMENTS

With my somewhat extended stay in graduate school, there are perhaps more people to thank and acknowledge than usual. I will nevertheless try to brief as possible, secure in the knowledge that it is merely an illusion of grandeur to think others will be offended at being left out from the acknowledgements section of a dissertation.

Beginning with my professors from undergraduate days who fanned the flames of my linguistic interests: Sally McLendon (Hunter College), Ricardo Ma. Nolasco (U.P.) and Viveca Hernandez (U.P.) are all gratefully acknowledged here.

My committee in Cornell: John Wolff, who has been a great friend, advisor and Vater-Doktor. I have learned immeasurably from his knowledge of all things Austronesian and I hope we are able to continue collaborating far into the future. John Whitman is a model of academic excellence, a bottomless font of knowledge, and an all around *übermensch* as is plain to all who know him. Draga Zec was always generous in lending me her expertise on all things prosodic and often provided much needed encouragement in my early days.

Of course, nobody survives graduate school without a loyal cadre of friends and in this regard I feel especially lucky. My first years were largely shaped by the inimitable friendship and camaraderie of the Fragile Four (Tanya Matthews, Diego de Acosta, Andrew Joseph + myself). Those days of dining on black beans and mayonnaise in the hermetically sealed “Man Hotel” would surely have been even bleaker without them. Anastasia (Bu) Riehl has played an enormously helpful role in my development and general mental health and happily remains to this day within my innermost circle. Another heroic Iowan, Edith Aldridge has also been a great friend over the years not to mention her contributions to my education in her capacity as fellow Austronesianist and beer enthusiast. I like to think that I honed my syntactic

argumentation skills jousting with her on the finer points of Austronesian phrase structure in my early days and I'm glad that we're able to continue our friendly jousting till now. Among the younger generation, Effi "the Corinthian" Georgala has been a great friend and has consistently lent me a helping hand when I was away from Ithaca without so much as a peep of complaint. Tova Friedman generously lent me her couch and car (complete with chauffer) on several occasions thus making life that much easier. Joe Pittayaporn, Devon Strolovitch, Yumiko Ishi, Johanna Brugman, Noam Andrews, Mie Hiramoto (Univ. of Hawaii) and the Cornell Filipino Association also deserve honorable mention here. Thess Savella deserves singling out as she blazed the way for me from Diliman to Ithaca. She helped me dissect Tagalog grammar on numerous occasions. I thank her for all of her patience and salute her dedication to her mother tongue.

I thank Angie and Shelia for putting up with one of the worst of their wards when it came to paperwork. Their administrative assistance saved me many a time.

Outside of Cornell, I owe Loren Billings for first getting me interested in clitics. Due to our collaboration, I came to develop many of the ideas for this thesis.

Austronesianists: Mark Donohue, Paul Kroeger, Bill Foley, Malcolm Ross and Bob Blust have all contributed immensely to my knowledge of language and linguistics.

The AFLA gang: Norvin Richards, Ileana Paul, Lisa Travis, Sandy Chung, Raph Mercado, Matt Pearson *inter alia* also deserve mention as providing a great atmosphere for the exchange of ideas over the years.

Junko Ito, Armin Mester, Jaye Padgett and Sandy Chung all helped to make my stint at UCSC an intellectually fruitful experience. Cathryn Donohue, my office mate, among several others who helped make it such a fun time, as well.

I cannot imagine what life would have been like without the generous support of the National Science Foundation Graduate Fellowship Program. The NSF allowed



me to escape teaching for several years and expand the breadth and depth of my investigations. I aspire to make good on their investment.

My family and friends in NYC: I thank my parents for their support and for never interfering in my strange plans despite seeming so irrational at the time. Their support and confidence has been invaluable.

(Tita) Bebot Galvan has helped me perhaps more than any other single source in getting to the bottom of some highly subtle patterns via a seemingly endless array of questions. I thank her for her help, patience and, most of all, friendship.

Jon Zachary, through his incessant chiding over the decades, has forced me to constantly reevaluate my every move, thereby strengthening my resolve.

My wife Amalia, who thought I was on the edge of the cusp of being finished with this when we got married 1.5 years ago. Had she only known. Her daily encouragement and vigilance kept me on track in the final stretch.

In sum, thanks to everyone, mentioned and un-, for all of your support over the years!

## TABLE OF CONTENTS

Biographical Sketch.....	iii
Dedication.....	iv
Acknowledgements.....	v
List of figures.....	x
List of tables.....	xi
List of tableau.....	xii
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1 Background.....	1
1.2 Tagalog.....	4
1.3 Some basic clitic facts.....	8
1.4 Organization.....	14
<b>Chapter 2: Approaches to clitic phenomena and properties</b> .....	<b>17</b>
2.1 Introduction.....	17
2.2 Clitics as morphological entities.....	18
2.2.1 The Zwicky criteria.....	18
2.2.2 Cardinaletti & Starke (1994/1999).....	22
2.2.3 Kiparsky & Condoravdi (2001,2004).....	39
2.3 Theories of second position: 2P(roblems).....	47
2.3.1 Weak phonology.....	50
2.3.2 Weak syntax and prosodic inversion.....	53
2.3.3 Distributed Morphology and morphological merger.....	59
2.3.4 Attachment to phonological hosts and Prosodic Subcategorization... ..	65
2.3.5 Optimality Theory approaches to 2P.....	76
2.4 An improved approach: 2P clitics as the output of Merge [F].....	83
2.4.1 The morphological framework.....	83
2.4.2 The phonological framework.....	93
2.4.3 Deriving 2P.....	94
2.4.4 Cluster internal ordering.....	101
2.5 Conclusion.....	106
<b>Chapter 3: The typology of clitic placement</b> .....	<b>108</b>
3.1 Introduction.....	108
3.1.1 The Klavans typology.....	107
3.2 Clitic types.....	111
3.2.1 Type 1: Initial, before, enclitic.....	111
3.2.2 Type 2: initial, before, proclitic.....	113
3.2.3 Type 3: initial, after, enclitic.....	117
3.2.4 Type 4: initial, after, proclitic.....	120
3.2.5 Type 5: final, before, enclitic.....	121
3.2.6 Type 6: final, before, proclitic.....	122
3.2.7 Type 7: final, after, enclitic.....	122
3.2.8 Type 8: final, after, proclitic.....	125
3.3 Divergent patterns of sisters and non-sisters: a new typology.....	125
3.3.1 Clitic typology and previous approaches.....	129

3.3.2 Prolegomena to a new account of clitic typology.....	135
3.4 Conclusion.....	139
<b>Chapter 4: Two types of structure building</b> .....	140
4.1 Introduction.....	140
4.2 Feature adjunction.....	142
4.2.1 The status of C <sup>0</sup> and the Phase Correspondence Principle.....	147
4.3 Concatenated clitics.....	155
4.3.1 Deriving enclitic prepositions.....	154
4.4 Head clitics and Feature clitics in Kwakw'ala.....	157
4.5 The factorial typology of adjoined clitics.....	163
4.6 Further consequences.....	167
4.6.1 Verb adjacent clitics and affixes.....	167
4.6.2 2P versus free pronominals.....	175
4.7 Conclusion.....	181
<b>Chapter 5: Tagalog clitic syntax and the clitic visibility condition</b> .....	183
5.1 Introduction.....	183
5.2 Clitic positioning within the clause.....	184
5.2.1 The architecture of the left periphery.....	184
5.3. Impenetrability and the Clitic Visibility Condition.....	191
5.3.1 Topics.....	197
5.3.2 Topicalized interrogatives and “clitic third”.....	201
5.3.3 Parentheticals.....	202
5.4 Clitic climbing.....	204
5.5 Clitic positioning within DP and non-verbal predicates.....	209
5.5.1 Nominal constituents.....	209
5.5.1.1 Modified nominal constituents.....	209
5.5.1.2 The left edge of the nominal clitic domain.....	214
5.5.1.3 Referentiality and impenetrability.....	220
5.5.2 Property predicates.....	228
5.5.3 Prepositional predicates.....	231
5.5.3.1 Complex prepositional predicates.....	235
5.6 Coordination.....	238
5.7 Ellipsis.....	243
5.8 Compounding, reduplication and clitic placement.....	251
5.9 Alternative theories of Tagalog clitic syntax.....	256
5.9.1 Kroeger 1993: the first daughter approach.....	256
5.9.2 Billings & Konopasky: verb adjacency.....	265
5.9.3 Balkan DP-internal clitics in Distributed Morphology.....	268
5.10 Conclusion.....	277
<b>Chapter 6: Conclusion</b> .....	278
6.1 Summary and prospects.....	278
<b>References</b> .....	280

## LIST OF FIGURES

Figure 1.1 The place of Tagalog in Austronesian.....	5
Figure 2.1 DM architecture .....	61
Figure 5.1. DM architecture (repeated) .....	271

## LIST OF TABLES

Table 1.1 Tagalog case markers.....	6
Table 1.2 Tagalog voice paradigms.....	6
Table 1.3 Tagalog pronominals.....	8
Table 1.4 Tagalog adverbial clitics.....	9
Table 2.1. Summary of coordination facts.....	30
Table 2.2. Summary of data on coordination, inclusive and linking facts.....	38
Table 2.3. Condoravdi & Kiparsky’s (2001) clitic typology.....	40
Table 2.4. Chamorro Pronouns.....	66
Table 2.5. Coordination facts.....	88
Table 2.6. Rankings for relative position of adverbial and pronominal clitics.....	103
Table 3.1. Klavans (1980, 1985) 8-way typology.....	110
Table 3.2. Tondano Pronominals.....	115
Table 3.3. Summary of attested clitic types.....	126
Table 3.4. Sister clitic typology.....	126
Table 3.5. Attested types of sister versus non-sister clitics.....	127
Table 4.1. K <sup>w</sup> ak <sup>w</sup> ala DP demonstrative clitics (unpossessed).....	160
Table 4.2. Factorial typology of concatenated and adjoined clitic interactions...	166
Table 5.1. Tagalog demonstratives.....	216
Table 5.2. Post-predicate head versus post-oblique phrase positioning of clitics...	260

## LIST OF TABLEAU

Tableau 2.1. Infix with C-initial stem (Prince & Smolensky 1993).....	78
Tableau 2.2. Infix with V-initial stem (Prince & Smolensky 1993).....	78
Tableau 2.3. 2P clisis (Anderson 2000, 2005).....	79
Tableau 2.4. Cluster-internal ordering of clitics.....	80
Tableau 2.5. The prosodic basis for non-initiality.....	99
Tableau 2.6. Cluster-internal ordering, simple monosyllabic case.....	104
Tableau 2.7. Cluster-internal ordering, simple monosyllabic case.....	104
Tableau 2.8. Cluster-internal ordering with mixed clitics.....	105
Tableau 4.1. Violation profile for domain initial adjoined morphemes.....	145
Tableau 4.2. Tagalog basic case.....	146
Tableau 4.3. Tondano basic case.....	147
Tableau 4.4. ALIGN (morph) >> ALIGN (XP,PPh) >> *WEAKSTART.....	164
Tableau 4.5. ALIGN (XP,PPh) >> ALIGN (morph) >> *WEAKSTART.....	165
Tableau 4.6. *WEAKSTART >> ALIGN (XP,PPh) >> ALIGN (morph).....	165
Tableau 4.7. ALIGN (morph) >> *WEAKSTART >> ALIGN (XP,PPh).....	166
Tableau 4.8. Tobler-Mussafia enclitics following a conjunction.....	171
Tableau 4.9. Tobler-Mussafia proclitics following a conjunction.....	171
Tableau 4.10. Tobler-Mussafia proclitics following Neg.....	171

# CHAPTER 1: INTRODUCTION

## 1.1 Background

Over a century ago, the great Indo-Europeanist Jacob Wackernagel formulated a generalization over the placement of a class of words in Indo-European languages and offered a historical explanation for their patterning. Wackernagel (1898) claimed that these elements were placed after the first word of the clause as a result of losing their accent. Second position, or “Wackernagel position”, as it came to be known, was seen as inherently weak and therefore appropriate for deaccented elements. Since Wackernagel’s seminal work, accounting for second position has been a challenge taken up by linguists of all stripes and theoretical persuasions. The challenge persists due to its relatively widespread distribution among the languages of the world and the variation found throughout its instantiations. We review here some of the problems posed by second position phenomena to linguistic theorizing.

The most obvious feature of second position (henceforth 2P) elements is that they involve a lack of semantic compositionality in the sense that they have no surface relationship with their semantic complements. This can be seen by comparing the position of free elements in Tagalog with their 2P counterparts. In (1), the free speaker-oriented adverb *malaman* ‘probably’ is positioned at the left edge of the sentence outside of negation, from where it takes scope.<sup>1</sup> Similarly, the genitive and nominative case phrases in the sentence are placed in their argument position after the predicate with which they compose. When we replace the free adverb with a clitic adverb of very similar meaning, and the full NP arguments with pronominals, as in (2), we find that the previously present compositionality breaks down. Because of the

---

<sup>1</sup> That (free) adverbs are positioned according to their scope is argued by Ernst (2002) and Kaufman (2005) for Tagalog. Compare English, ‘He probably doesn’t eat meat.’ with #‘He doesn’t probably eat meat.’ The second sentence is only possible under the reading where ‘probably’ scopes under negation, i.e., ‘He doesn’t probably eat meat, he definitely eats meat.’

2P requirement of these elements in Tagalog, the adverb is internal to negation and both pronominal arguments are separated from the predicate by a locative adjunct. Explaining how and why this lack of compositionality is allowed so systematically can be considered the first challenge presented by 2P phenomena.

- (1) Malama<sub>1</sub> hindi dito ka~ka-usap-in na<sub>1</sub>=ma<sub>1</sub>na=pulis<sub>2</sub> a<sub>1</sub>=ma<sub>1</sub>a=preso<sub>3</sub>  
 probably NEG here CO~INCM-talk-PV GEN=PL=police NOM=PL=prisoner  
 ‘The police probably won’t talk to the prisoners here.’
- (2) Hindi=yatà<sub>1</sub>=nila<sub>2</sub>=sila<sub>3</sub> dito ka~ka-usap-in  
 NEG=PROB=3P.GEN=3P.NOM here CO~INCM-talk-PV  
 ‘They probably won’t talk to them here.’

As implied by its very name, the notion of “second position” additionally poses a basic challenge to the well accepted idea that the human language faculty never relies on counting (Chomsky 1965). To take Chomsky’s (1965:55) example, there are no morphemes in human language which require insertion in the mid-point of a relevant domain, e.g., after the fourth word in an eight word domain, as in (3)a, and after the fifth word in a ten word domain, as in (3)b.

- (3) a.  $x + [1-8] \Rightarrow *[1] [2] [3] [4] x [5] [6] [7] [8]$   
 b.  $x + [1-10] \Rightarrow *[1] [2] [3] [4] [5] x [6] [7] [8] [9] [10]$

More to the point, there are also no morphemes which follow the more plausible rule of insertion after a second unit in their domain, as shown in (4). This makes the existence and relative popularity of rules such as (5) all the more surprising; either the grammar does have a limited ability to count, or something else is at work here.



$$(4) \quad x + [1-8] \quad \Rightarrow \quad *[1] [2] x [3] [4] [5] [6] [7] [8]$$

$$(5) \quad x + [1-8] \quad \Rightarrow \quad [1] x [2] [3] [4] [5] [6] [7] [8]$$

One family of explanations sees 2P not so much as a syntactic target in and of itself but rather as the most local refuge for morphemes avoiding first position, an idea which can be traced to Wackernagel. The fact that 2P really does involve avoidance of initial position is supported by the fact that there are no convincing examples of 2P from the right edge, as in (6).

$$(6) \quad [1-8] + x \quad \Rightarrow \quad *[1] [2] [3] [4] [5] [6] [7] x [8]$$

This brings us to a third problem. If avoidance of first position is best characterized as a phonological issue, i.e., phonologically “weak” elements strive to stay out of prosodically “strong” positions, then there arises a question of modularity. The canonical domain of phonology encompasses such phenomena as segmental alternations, syllable structure, metrical prominence, among others, but it is not typically held to account for the order of morphemes and words. If phonology does play a role in the positioning of certain morphemes, we are led to ask what the extent of its power is in permuting morphological elements. Concretely, 2P phenomena militate against simplistic notions of the syntax-phonology interface in which, syntax, the ordering component, “turns off” and sends its output to the articulatory system. On the other hand, there is clearly a limit to what kind of phonological information can be relevant to ordering. Zwicky & Pullum (1986 *et seq*) argue, for instance, that our theory of grammar should not countenance a rule which moves words beginning with /b/ to clause initial position, as no rule of this sort is found in natural language.

In this dissertation, we seek to answer certain questions of the syntax-phonology interface by a detailed look at Tagalog 2P clitics. Tagalog clitics have

played a prominent role in the clitic literature since their discussion by Bloomfield (1917) and have been of central interest to much following work (Schachter & Otones (1972), Schachter (1973), Kroeger (1993, 1998), Sityar (1989), Anderson (1996, 2005), Billings & Konopasky (2004), Billings (2005), Richards (2004), Ramos (1971), Wolff, Centeno & Rau (1991), Kaisse 1981, 1982, Zwicky 1977, Halpern 1995). Yet, the Tagalog clitic data around which the above theorizing has taken place is severely lacking. Because it is only through examining complex environments that we can adjudicate between the many possible theoretical accounts available, the first goal of this dissertation is to expand the database. Here, through the use of electronic text searches, frequency data, and careful informant work, we discover several generalizations which have gone unnoticed in the literature. As I will show, these generalizations have far ranging implications for a more general theory of clitic placement.

## **1.2 Tagalog**

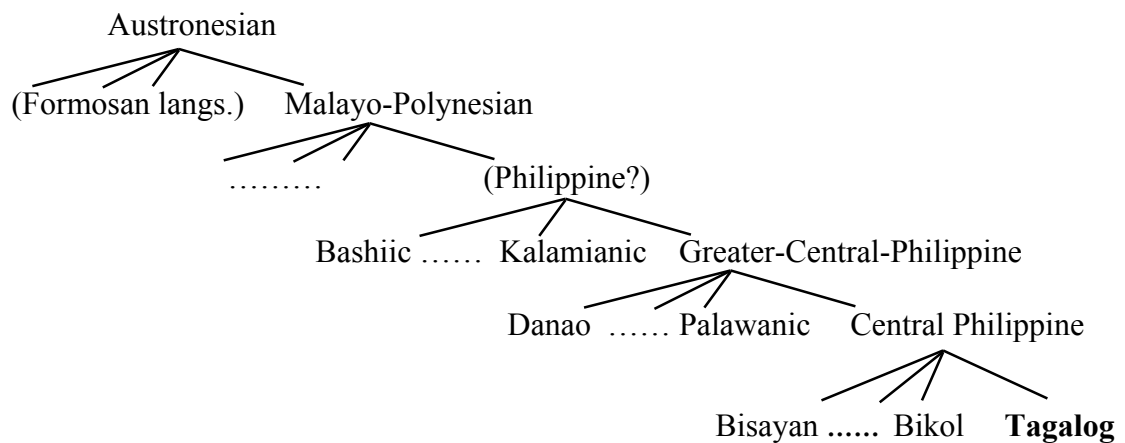
Tagalog is an Austronesian language spoken by over 15 million speakers in the Philippines and around the world.<sup>2</sup> On the island of Luzon, it is spoken natively in the provinces of Batangas, Bulakan, Rizal, Nueva Ecija, Laguna and Quezon and also spoken as a first language by speakers who are ethnically non-Tagalogs in Manila and other large cities. Outside of Luzon, it is spoken natively on Marinduque, Mindoro and Palawan, in the last two islands by descendents of relatively recent migrants from Luzon. It is now spoken by the majority of Filipinos to varying degrees of fluency in its capacity as the national language.<sup>3</sup>

---

<sup>2</sup> For a good overview of the Austronesian family see Adelaar & Himmelmann (2005).

<sup>3</sup> In this capacity Tagalog is known as Filipino. Ethnic sensitivities required the use of a more neutral name in the place of an ethnonym to refer to the national language although no actual differences exist between the two.

Although much work remains to be done on the subgrouping of Philippine languages amongst themselves and the position of Philippine languages in relation to Malayo-Polynesian, current classification places Tagalog as shown in Figure 1.1<sup>4</sup> The basis for a Philippine subgroup has been debated (Reid 1979, 1981, 1982, Zorc 1986 Blust 2005). Blust (1991) presents evidence for a group he calls Greater Central Philippine while Zorc (1986) discusses evidence for the smaller Central Philippine group.



**Figure 1.1.** The place of Tagalog within the Austronesian family tree

Typologically, Tagalog is a robustly predicate initial language which allows scrambling of phrases in the post-predicate position. Content words of any notional category may head the predicate phrase without the use of copula, as seen in (7).

- (7) a. Gúro aᅇ=babae  
 teacher NOM=woman  
 ‘The woman is a teacher.’
- b. Nag-túro aᅇ=babae  
 AV.BEG-teach NOM=woman  
 ‘The woman is teaching’
- c. Nása=labas aᅇ=babae  
 OBL=outside NOM=woman  
 ‘The woman is outside’

<sup>4</sup> The Meso-Philippine subgroup is not as well-supported as the others. Central Philippine is more secure although much work remains to be done here, too.

All arguments are case marked by one of three case markers, shown in Table 1.1. Names of people are marked by a special set of personal case markers which have their own plural marking. With the impersonal case markers, plurality is indicated by the more general plural marker *maṅa*.<sup>5</sup>

**Table 1.1.** Tagalog case markers

Impersonal		Personal	
sg.	pl.	sg.	pl.
<i>aṅ</i>	<i>aṅ maṅa</i>	<i>si</i>	<i>sina</i>
<i>naṅ</i>	<i>naṅ maṅa</i>	<i>ni</i>	<i>nina</i>
<i>sa</i>	<i>sa maṅa</i>	<i>kay</i>	<i>kina</i>

Most discussion of Tagalog in recent literature revolves around the distinctive alignment pattern found with so-called “focus type languages” in the Philippines (see Kaufman 2009 and references therein). These languages are distinctive because of their rich voice system which allows turning almost any argument or adjunct into the subject of the sentence. Whereas many other languages allow the promotion of different sorts of objects to subject through the use of applicatives and passives, Philippine languages are interesting in not treating any single voice as morphologically basic. All voices appear equally derived from their roots. The primary voices are shown in Table 1.2. Two voices, the actor voice and the patient voice, are additionally shown in the abilitative mode.

**Table 1.2.** Tagalog voice paradigms

	Patient voice (PV)	Locative voice (LV)	Conveyance voice (CV)	Actor voice I (AV)
Root	sábi ‘say’	sábi ‘say’	bigay ‘give’	bigay ‘give’
Infinitive	sabí-hin	sabí-han	i-bigay	b<um>igay
Perfective	s<in>ábi-Ø	s<in>abí-han	i-b<in>igay	b<um-Ø>igay
Progressive	s<in>á~sábi-Ø	s<in>á~sábi-han	i-b<in>í~bigay	b<um-Ø>í~bigay
Prospective	sá~sabí-hin	sá~sabí-han	i-bí~bigay	Ø-bí~bigay

<sup>5</sup> Unlike other Southeast and East Asian languages, the lack of plural marking in Tagalog and other Philippine languages is deterministic. Arguments not marked for plurality can only be interpreted as singular or generic but cannot indicate plural referents.

**Table 1.2.** (Continued)

	Actor voice II (AV)	Actor voice abilitative (AV.ABL)	Patient voice abilitative (PV.ABL)
Root	bigay ‘give’	kítà ‘see’	kítà ‘see’
Infinitive	mag-bigay	maka-kítà	ma-kítà
Perfective	nag-bigay	naka-kítà	na-kítà
Progressive	nag-bí~bigay	naká~ka-kítà	na~kí-kítà
Prospective	mag-bí~bigay	maká~ka-kítà	ma~kí-kítà

The morphological atoms of the above paradigms are not easy to tease out. There are three null morphs which are indicated above where they are expected on the basis of their paradigms: the patient voice in the BEGUN (perfective and progressive) aspects, the BEGUN aspect in the actor voice, and the actor voice in the prospective aspect. The last two instances of null morphology have overt counterparts in other Philippine languages but the null marking of the patient voice in the presence of the <in> infix is an ancient pattern which can be reconstructed to Proto-Austronesian.

The aspectual system is analyzed here by decomposing the three aspects into two features [ $\pm$ BEGUN] and [ $\pm$ INCOMPLETE]. With this feature system, the two aspect marking morphemes CV reduplication and the infix <in> can now be interpreted as marking [+INCOMPLETE] and [+BEGUN], respectively. The combination of these two features yields the three aspects shown in (8). Thus, when we refer in the text to perfective aspect, this will be understood to be indicated by the begun affix, and so forth.

- (8) a. BEGUN > Perfective  
 b. BEGUN + INCOMPLETE > Progressive  
 c. INCOMPLETE > Prospective

This dissertation remains neutral regarding the alignment system. Ergative analyses have been put forth by Gerdtz (1988), De Guzman (1988), Aldridge (2004) and Liao (2005) among others and accusative analyses have been argued for by

Kroeger (1991/1993), Richards (2000), Rackowski (2002), Rackowski & Richards (2005). Others yet have argued for a more symmetrical approach to the voice system setting it apart from canonical accusative and ergative systems (Naylor 1980, Foley 1998, 2008, Kaufman 2009).

### 1.3 Some basic clitic facts

Tagalog clitics fall into two broad classes: pronominal and adverbial. These are given in Table 1.3 and Table 1.4, respectively.

**Table 1.3.** Tagalog pronominals

Trad. labels	Gloss	Features	NOM	GEN	NOM	GEN	OBL
			CLITIC		FREE		
1 <sup>st</sup> sing.	1S	[1]	= <i>ako</i>	= <i>ko</i>	<i>ako</i>	<i>ákin</i>	<i>sa ákin</i>
2 <sup>nd</sup> sing.	2S	[2]	= <i>ka</i>	= <i>mo</i>	<i>ikaw</i>	<i>iyo</i>	<i>sa iyo</i>
3 <sup>rd</sup> sing.	3S	[∅]	= <i>siya</i>	= <i>niya</i>	<i>siya</i>	<i>kaniya</i>	<i>sa kaniya</i>
1 <sup>st</sup> excl. pl.	1+3	[1, p]	= <i>kami</i>	= <i>námin</i>	<i>kami</i>	<i>ámin</i>	<i>sa ámin</i>
(1 <sup>st</sup> dual)	1+2	[1,2]	= <i>kata/kita</i>	= <i>ta</i>	<i>kata/kita</i>	<i>kanita</i>	<i>sa kanita</i>
1 <sup>st</sup> incl. pl.	1+2P	[1,2,p]	= <i>táyo</i>	= <i>nátin</i>	<i>táyo</i>	<i>átin</i>	<i>sa átin</i>
2 <sup>nd</sup> pl.	2P	[2,p]	= <i>kayo</i>	= <i>ninyo</i>	<i>kayo</i>	<i>inyo</i>	<i>sa inyo</i>
3 <sup>rd</sup> pl.	3P	[∅,p]	= <i>silá</i>	= <i>nila</i>	<i>silá</i>	<i>kanila</i>	<i>sa kanila</i>
			Portmanteau forms: [1.GEN+2.NOM] = <i>kita, kita</i>				

**Table 1.4.** Tagalog adverbial clitics

	CLITIC	FREE
aspect	= <i>na</i> ‘already’	∅
	= <i>pa</i> ‘still’	∅
focus	= <i>din</i> ‘also’	∅
	= <i>man</i> ‘even’	∅
	= <i>naman</i> ‘switch topic’	( <i>naman</i> )
	= <i>ηà</i> ‘emphasis’	∅
	= <i>lan</i> ‘only’	∅
	= <i>láman</i>	<i>láman</i>
	= <i>talaga</i> ‘emphasis’	<i>talaga</i>
politeness	= <i>pò</i> , = <i>hò</i> ‘politeness’	∅
mood	= <i>pala</i> ‘surprise’	∅
	= <i>yátà</i> ‘perhaps’	∅
	= <i>sána</i> ‘hopefully’	<i>sána</i>
	= <i>náwa</i> ‘hopefully’	<i>náwa</i>
	= <i>ba</i> ‘question marker’	∅
	(= <i>baga</i> )	( <i>baga</i> )
	= <i>daw</i> reported speech	∅

As seen in Table 1.3, there exist both free and clitic sets of nominative and genitive pronouns. As we will see shortly, these are differentiated only by their position in the sentence and their focusability. Oblique pronouns have no clitic counterparts although in less formal registers they may be positioned as clitics.<sup>6</sup>

Many of the adverbial clitics, shown in Table 1.4, have no free counterparts although others do have homophonous free variants which can be topicalized. While the absence of free forms can be explained for many of the adverbials by their monosyllabic composition, this would not account for lack of free counterparts for *pala* and *yátà*. The adverbial *naman* has a free counterpart only in the most colloquial style while *baga* is a dialectal form, absent in Manila Tagalog.

The forms in Tables 1.3 and 1.4 will serve as the primary material with which we investigate theories of clitic morphology and syntax.

<sup>6</sup> Unfortunately, I will not be able to address the phenomenon of optional clisis in this dissertation beyond a few very general remarks. Optional clisis applies to demonstratives, oblique pronouns and proper names. For a more detailed discussion see Billings (2005).

Returning to clitic syntax, we see in sentences (9)-(14) that the pronominal clitic appears directly after the first word (to be defined below).<sup>7,8</sup>

- (9) Na-túto=**siya** naŋ=wíka=ŋ Instsik  
 AV.BEG-learn=3S.NOM GEN=language=LNK Chinese  
 ‘She learned Chinese’
- (10) Hindí=**siya** na-túto naŋ=wíka=ŋ Instsik  
 NEG=3S.NOM AV.BEG-learn GEN=language=LNK Chinese  
 ‘She didn’t learn Chinese’
- (11) Saan=**siya** na-túto naŋ=wíka=ŋ Instsik?  
 NEG=3S.NOM AV.BEG-learn GEN=language=LNK Chinese  
 ‘Where did she learn Chinese?’
- (12) Kay=Yao=**siya** na-túto naŋ=wíka=ŋ Instsik  
 P.OBL=Yao=3S.NOM AV.BEG-learn GEN=language=LNK Chinese  
 ‘She learned Chinese from Yao.’
- (13) Dápat=**siya** na-túto naŋ=wíka=ŋ Instsik  
 NEG=3S.NOM AV-learn GEN=language=LNK Chinese  
 ‘She should have learned Chinese.’
- (14) Ma-bilis=**siya=ŋ** na-túto naŋ=wíka=ŋ Instsik  
 NEG=3S.NOM=LNK AV.BEG-learn GEN=language=LNK Chinese  
 ‘She learned Chinese quickly.’

This position is not available to non-clitic arguments. If we replace the subject clitics in (9)-(14) with a full NP in pre-predicate position, as in (15), the results are consistently ungrammatical.

- (15) Hindí [**\*aŋ=paŋúlo**] na-túto naŋ=wíka=ŋ Instsik [**aŋ=paŋúlo**]  
 NEG NOM=president AV.BEG-learn GEN=language=LNK Chinese NOM=pres.  
 ‘The president didn’t learn Chinese’

<sup>7</sup> In (9)-(12), this is the only position available for clitics. In (13)-(14), other options exist (see chap. 5).

<sup>8</sup> Second position elements will be referred to as CLITICS regardless of whether these items actually display any phonological dependency. The element preceding the clitics in all these cases will be referred to as the CLITIC HOST. The relationship between clitic and host is indicated with the equals sign. For the case markers such *aŋ*, *naŋ*, *sa*, this will indicate a rightwards looking dependency and for second position clitics, this will indicate a leftwards looking dependency.



It is already apparent from the simple data above that clitics surface as close as possible to the left edge of the clause without actually appearing on the edge. But, as noted above, an interesting fact about 2P clitics is that both the prosodic and syntactic properties of the domain initial constituent effects their positioning. For instance, in (12), the clitic doesn't follow the first morphological word but rather the second one. The morphological word *kay* is a monosyllabic function word which functions to case mark its complement *Yao*. Several characteristics of such words have been independently recruited to explain their non-host behavior in other languages.<sup>9</sup> It is clear from data such as (16) and (17) that monosyllabic function words cannot host clitics in Tagalog and this falls into line with cross-linguistic expectations that only prosodic words can host clitics.

(16) Kuŋ=hindi=**ka** maka-ratiŋ...  
 if=NEG=2S.NOM AV.ABL-arrive  
 'If you can't arrive...'

(17) Sa=báhay=**ka**  
 OBL=house=2S.NOM  
 'You're at the house.'

However, it turns out that the conditions on hosting clitics in Tagalog are more complex than the requirement of prosodic wordhood. Specifically, there also exist phrases which cannot be intruded upon by certain clitics for structural reasons. For instance, as first noted by Schachter & Otnes (1972), focus fronted oblique phrases cannot be intruded upon by pronominal clitics, as shown in (18).

---

<sup>9</sup> Inkelas (1990) and Zec & Inkelas (1990) originally argued that clitic hosts must be prosodic word while Marantz (1988) suggested that the relationship between certain heads and their complements cannot be interrupted by clitics. More syntactically oriented accounts such as Cavar (1996) *inter alia* connect this to the fact that non-hosts cannot be stranded by regular syntactic operations.

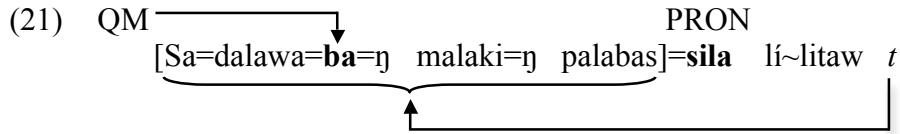
- (18) [Sa=dalawa=η malaki=η palabas]=**sila** lí~litaw  
 OBL=two=LNK big=LNK show=3P.NOM AV.INCM~appear  
 ‘They will appear in two big shows.’

The fact that the material within the bracketed phrase in (18) cannot host the pronominal will be referred to here as impenetrability, and forms one of the central topics of the thesis. Interestingly, impenetrability does not appear to effect all clitics but only pronominal ones. For instance, compare the position of the pronominal clitic with the position of the question marking clitic in (19) and (20).

- (19) [Sa=dalawa=**ba**=η malaki=η palabas]=**sila** lí~litaw?  
 OBL=two=QM=LNK big=LNK show=3P.NOM AV.INCM~appear  
 ‘Will they appear in two big shows?’
- (20) [Saan=**pa**=**ba**=η panadéro]=**kayo** b<um>í~bili nanη=tinápáy?  
 where=still=QM=LNK baker=2P.NOM <AV.BEG>INCM~buy GEN=bread  
 ‘From which other baker do you buy bread?’

We find the adverbial clitics in precisely the position we expect if clitics follow the first prosodic word within their domain (roughly, the clause). So whatever accounts for the impenetrability to pronominal clitics cannot be an absolute constraint otherwise it should effect all clitics equally. In Halpern’s (1995) theory of clitic placement, 2P clitics are generated as regular syntactic items in the phrase structure and can have phrasal hosts when a phrase happens to move around them. If, by the end of the syntactic derivation, nothing precedes a clitic, the clitic will undergo inversion with a following morphological or syntactic constituent (i.e. a word or phrase). Because not all clitics need be generated in the same position, the following analysis presents itself on such a theory. The oblique phrase moves to a peripheral position for focus reasons and this position happens to lie in between the underlying position of the question marker and the pronominals, as shown in (21). Because syntax does not provide a host preceding the question marker, Prosodic Inversion kicks in and inverts the hostless clitic with the following prosodic word. Consequently, the question

marker will follow *dalawa* but the pronominal will be able to remain in its base position because the syntax has already provided its host.



This account can handle the above data elegantly but, unfortunately, makes the wrong predictions for environments in which there is another host further to the left. Given the derivation sketched in (21), we do not expect that pronominals would continue their leftward migration if they are generated below the focus phrase. But as seen in (22), pronominals typically do appear before the focus phrase given an appropriate host.

- (22) Hindi=**ba**=**sila** [sa=dalawa= $\eta$  malaki= $\eta$  palabas] lí~litaw  
 NEG=QM=3P.NOM OBL=two=LNK big=LNK show AV.INCM~appear  
 ‘Won’t they appear in two big shows?’

This suggests that the problem is better framed in terms of impenetrability rather than clitic domains. I show that there must exist a surface relation between 2P pronominal clitics and the predicates which assign them thematic roles, as formulated in (23).

- (23) Clitic Visibility Condition (CVC)  
 For every argument clitic  $\alpha$  assigned a thematic role by a predicate head  $\beta$ , the minimal maximal projection *linearly containing*  $\alpha$  must dominate  $\beta$ .

Essentially, the CVC prevents pronominal clitics from being separated from their predicates by being embedded within higher phrases. This does not apply to adverbial clitics because adverbials have no theta dependency and indeed have no dependencies at all to overt material within the clause. Their position is thus freer and can be handled entirely by the prosodic component.

The fact that clitic placement is basically prosodic (after the first prosodic word) but displays a syntactic filter in the form of the CVC poses a challenge to most theories of clisis and indeed many of the serial grammatical architectures posited in the literature which give phonology “last pass” on syntactic derivations. Overall, this supports the copresence model of the syntax-phonology interface (Zec & Inkelas 1990), where output structures must satisfy constraints from different components simultaneously.

This dissertation thus aims to offer a more typologically responsible theory of clisis by deriving only attested clitic patterns and by making the crucial distinction between feature clitics and syntactic head clitics. Pronominal clitics are adjoined by an operation which adds syntactic features to preexisting nodes without building syntactic material. This derives their penchant for movement in comparison to syntactic head clitics which are not found to be misaligned from their hosts. Finally, the syntactic visibility condition shows that the outlines of syntactic structure must still be present at the point when clitic position is determined. The fact that clitic hosts cannot be derived by the regular syntax, however, shows that syntax only functions here as a filter and not as the primary positioning mechanism for clitics.

#### **1.4 Organization**

This dissertation is organized as follows. In chapter two I examine different notions of clitic in the literature and the extent to which they agree with Tagalog 2P elements. We find that theories of discrete clitic types do not offer an easy classification for the combination of properties shown by Tagalog clitics and that there exists a major paradox in the syntactic behavior Tagalog 2P clitics and their coordinatability. We then review several theories of clitic placement, looking briefly at syntactic theories and focusing on phonology-oriented theories. Although phonology oriented theories appear to work best for Tagalog there are several

important issues that these proposals fail to resolve. Specifically, phonological theories are hard put to handle syntactic impenetrability phenomena, to be explained below, and Tagalog clitics' lack of phonological dependency. Finally, I offer a prosody based theory of non-initiality couched in Optimality Theory which allows us to derive second position without prosodic dependency.

Stepping back from Tagalog, the third chapter examines clitic typology from a crosslinguistic perspective. Klavans (1989) argues for a very wide-ranging typology of clitics which were positioned and parsed prosodically on the basis of three independent parameters. Here, I reexamine this typology by dividing clitics into two classes: syntactic heads (more generally morphosyntactic sisters) and feature bundles. In the former group we find elements such as adpositions, case markers and complementizers while in the latter group we find elements such as argument clitics and certain adverbial clitics. These two groups behave quite differently from each other with regard to their positioning and prosodic parsing. It is only when we accept this basic division that we can make sense of the clitic typology on an explanatory level.

In chapter four, I argue that the difference between feature clitics and syntactic heads is derived from the distinct ways in which heads and features are introduced into the syntax. Heads are concatenated with their complements via the operation Merge Terminal whereas feature bundles are adjoined to preexisting nodes via Merge Feature and then subject to late spell-out. As a result, syntactic heads are “protected” by the phrase structure which dominates them in several ways. When prosodic phrases are aligned to syntactic phrases, these head clitics will be naturally included in the relevant prosodic structure. On the other hand, because adjoined features are not taken to be associated with a terminal node in the syntax, they are left out when syntactic phrases are mapped to prosodic phrases. In Optimality Theory, this does not ensure that they will not be parsed, as there is always the possibility that a prosodic phrase



## CHAPTER 2: APPROACHES TO CLITIC PROPERTIES AND PHENOMENA

### 2.1 Introduction

In this chapter we review properties of clitics and various theoretical approaches to clitic phenomena in the literature. We are primarily interested here in defining clitics within a typology of morphological elements and exploring different theories for their placement, and to a lesser extent, their integration into prosodic structure. We focus here especially on theories of second position and their application to Tagalog.

Many proposals have been made to include clitics as a discrete intermediate category in between word and affix, in accordance with the results of certain phonological and syntactic diagnostics. Unfortunately, there is little agreement between the different theories which treat clitics as discrete universal categories. In §2.2, we apply various syntactic diagnostics to Tagalog, and other Austronesian clitics, with the aim of evaluating competing proposals regarding the nature of these categories. In §2.3, we examine several current theories of 2P clisis, showing the inadequacy of a primarily syntactic account in capturing the facts of Austronesian 2P patterns. The approaches most applicable to our data are Halpern's (1995) theory of Prosodic Inversion and Inkelas' (1990) theory of Prosodic Subcategorization. A careful comparison with the Tagalog facts, however, will show that neither of these approaches can handle the central data in a satisfactory manner. The Optimality Theoretic treatments of 2P phenomena developed by Anderson and Legendre are also discussed in this section. This general approach, modified to handle prosodic word clitics more naturally, is adopted in the following chapters. Finally, I summarize the findings and conclude in §2.4.

## 2.2 Clitics as morphological entities

### 2.2.1 *The Zwicky Criteria*

Beginning with Zwicky (1977, 1985) Zwicky & Pullum (1983), a number of syntactic tests have been used to examine the status of elements which are intermediate between affixes and full words. Although a near consensus has developed in the more careful typological literature<sup>1</sup> that a *cross-linguistic* definition of clitic cannot be consistently upheld on the basis of such tests (Zwicky 1994, Sadock 1995), these tests are nonetheless crucial in pointing us towards the locus of variation in the gray area between affixes and words. It goes without saying that the theoretical significance of these diagnostics has been understood in different ways by different researchers. Some have approached the gray area as a complete cline upon which no discrete categories can be identified (Givón 1971, Janse 1998) while others have posited various numbers of discrete clitic categories, each showing different levels of prosodic and syntactic attachment to their host. Among this latter school, Halpern & Fontana (1994) have posited a two way split (between  $X^0$  and  $X^{\max}$  type clitics) while Cardinaletti & Starke (1994/1999) and Kiparsky & Condoravdi (2001, 2004) have proposed a more fine grained three way split. We will first investigate some common characteristics of Philippine clitics, using Tagalog to illustrate, and then attempt to answer whether or not any of the above discrete category theories have the power to explain these properties.

In one of the most influential modern works on clitics, Zwicky (1977) proposes a rough descriptive typology with three categories: “simple clitic”, “special clitic” and “bound word”. These do not correlate with particular syntactic structures

---

<sup>1</sup> See for instance, the papers in Dixon & Aikhenvald (2006), and Nevis (1986) who argues against the notion of clitic as a grammatical category). Zwicky (1994) himself, who was largely responsible for the notion of a discrete universal class of clitics (regardless of whether this was actually his position), states, “clitics are unlikely [to] constitute a unified class for the purposes of theorizing about the nature of grammar.”



but rather correspond to roughly three types of behavior: simple clitics are free morphemes that are “phonologically subordinate to a neighboring word”, special clitics are unaccented, bound and alternate with a stressed free form of the same meaning and similar phonology, and bound words are clitic elements which have no corresponding full forms (Zwicky 1977:5). The theoretical status of the final category, was never entirely certain, with its main characteristic being the lack of a full form, and was consequently abandoned in later work (Zwicky & Pullum 1983).

The “Zwicky criteria” for clitics, as it has come to be known, is introduced below and examined in relation to Tagalog second position elements.

**(i) Clitics can exhibit a low degree of selection with respect to their host while affixes exhibit a high degree of selection with respect to their stems.**

Aspect is marked by both affixes and clitics in Philippine languages. In Tagalog, combinations of affixes and reduplication mark the perfective, progressive, and prospective aspects. The completive, roughly ‘already’, and incompletive, roughly ‘still’, on the other hand, are marked by the clitics =*na* and =*pa*, respectively. While the affixal markers must attach to the predicate head, the clitic markers typically attach to the first word within the clause, regardless of its category, as exemplified by (1).

- (1) Búkas=**pa**=ako                      Ø-**a**~alis  
tomorrow=**still**=1S.NOM AV-INCM~leave  
‘I’m leaving tomorrow yet.’

**(ii) Semantic idiosyncrasies are more characteristic of affixed words than of clitic groups.**

In (2), three Tagalog examples are found of idiosyncratic meanings which are obtained only with a particular affix or combination of affixes. No similar examples can be found with clitics.<sup>2</sup>

- (2) a. *bígát-in*                      b. *pag-tulúṅ-an*                      c. *i-bato*  
       heavy:NMLZ-PV                      TRNS-help-LV                      CV-stone  
       ‘big shot’                              ‘gang up on SUBJ’                      ‘throw’

**(iii) Syntactic rules do not affect clitic groups while they can affect affixed words.**

No movement rules can make reference to the clitic cluster nor to the host+clitic constituent in Tagalog. When an adjunct interrogative appears in a simple sentence, 2P clitics must follow it directly in the left periphery, as seen in (3)a. When long movement takes place, however, as in (3)b, 2P clitics are clause-bound and thus cannot move cyclically with the interrogative into the higher clause.

- (3) a. *Saan<sub>i</sub>[=*sila*]                      Ø-pu~punta[\*=*sila*] *t<sub>i</sub>* ?  
       where=3P.NOM                      AV-INCM~go  
       ‘Where are they going?’*
- b. *Saan<sub>i</sub>[\*=*sila*]                      s<in>ábi-Ø=ṅ                      *t<sub>i</sub>* Ø-pu~punta[=*sila*] *t<sub>i</sub>* ?  
       where=3P.NOM                      <BEG>say-PV=LNK                      AV-INCM~go  
       ‘Where were they said to be going?’*

**(iv) Clitics can attach to material already containing clitics, but affixes cannot.**

Morphological layering can be exemplified by the Tagalog example in (4). Affixes must precede the clitic cluster as in (4)a, and cannot follow, as in (4)b.

- (4) a. *Káin-in=mo=na=ṅà*    b. \**Káin=na-hin*  
       eat-PV=2S.GEN=ALRD=EMPH    eat=ALRD-PV  
       ‘Eat it already!’

<sup>2</sup> The only example which comes to mind is *ṅayon=din* now=also ‘immediately’, although this is apparently found in a large number of Austronesian languages (e.g., Malay/Indonesian *sekarang juga* now also) and may be semantically decomposable in some way.

(v) **Clitics cannot occur in complete isolation.**

This holds true for all monosyllabic clitics (=ka 2S.NOM, =na already, =pa still, =ba QM, =ηà EMPH, =laη ‘only’, =din ‘also’) and for most disyllabic adverbial clitics (yátà ‘perhaps’, pala SURPRISE, naman SWITCH TOPIC, lámaη ‘only’). Disyllabic nominative pronominals, however, in addition to certain disyllabic adverbial clitics can appear in isolation. These independent versions must be analyzed as homophones if we are to uphold this criterion.

(vi) **Clitics are strictly ordered with respect to adjacent morphemes while independent words may exhibit free ordering.**

The ordering of multiple clitics within the clitic cluster is strict overall, with only few areas of variability. The general pattern is as follows: 1σ pronouns > 1σ adverbials > 2σ adverbials > 2σ pronouns, in addition to a case constraint, GENITIVE > NOMINATIVE (to be discussed further below). The case ordering constraint is only tendency in the syntax of full NPs and the syllable count and adverb-argument layering constraints do not come into play at all in the regular syntax.

(vii) **Clitics follow simple principles governing their distribution while the combinatory possibilities of independent words are complex.**

This is a highly informal diagnostic without a reliable definition of “simple”. We can, however, point to several syntactic options which are available to DPs but not clitics. Scrambling of DPs in the post-predicate field, for instance, is common in Philippine languages. 2P clitics however, cannot be moved by such operations and must cluster together after the first legitimate host within their domain.

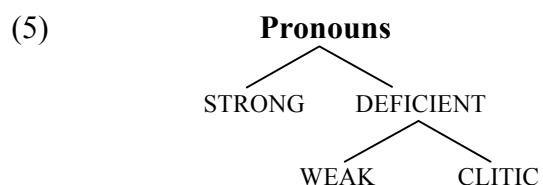
(viii) **Clitics are usually less morphologically complex than independent words.**

All post-predicate NP arguments in Tagalog must be marked by one of the three case markers, *ang*= NOM, *ng*= GEN or *sa*= OBL. Pronouns show the same case distinctions but only the oblique set (which is not comprised of bona-fide 2P clitics) is overtly marked by a case marker. The genitive and nominative clitics are inherently case marked and thus cannot be marked further by case markers. Adverb clitics are also morphologically simplex in Tagalog, as opposed to full word adverbials, which may be morphologically marked for a number of features.

The general diagnostics above all show Tagalog 2P clitics to be significantly different from both free words and affixes on several levels. In the next sections we look at whether more refined diagnostics locate Tagalog clitics together with those of other languages within a discrete linguistic category.

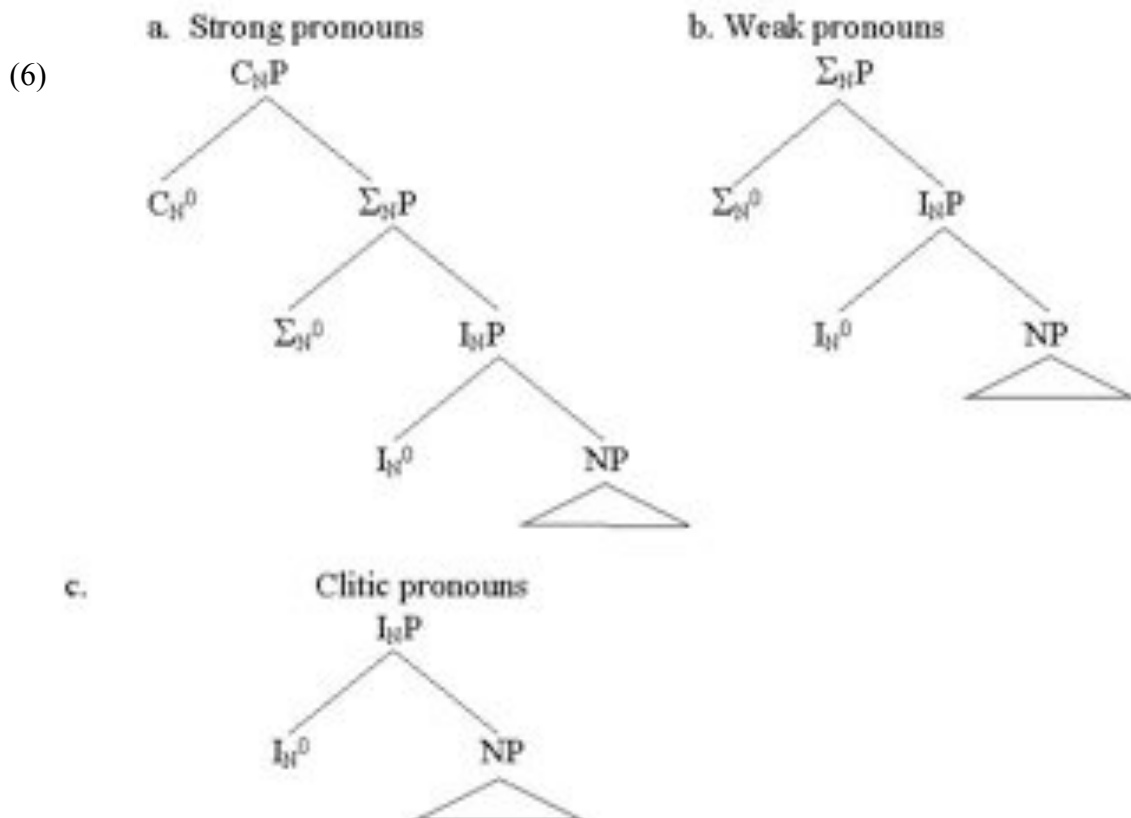
### 2.2.2 *Cardinaletti & Starke (1994/1999)*

In one of the earliest generative treatments of clitic phenomena, Kayne (1969, 1975) showed that French clitics differ from free words in several other syntactic respects, most importantly, these clitics were shown to disallow modification and conjunction. These observations were later formalized in a theory of pronominal weakness by Cardinaletti & Starke (1994/1999) (henceforth C&S) who proposed the tri-partite taxonomy shown in (5).



In C&S's theory, strong pronouns are equivalent to full noun phrases while deficient pronouns are characterized by reduced syntactic structure. This reduced structure also has morphological and phonological consequences, aside from the obvious syntactic differences. Deficient pronouns are divided into two classes: weak pronouns, which are XP elements, and clitic pronouns which are  $X^0$  elements.

Both deficient pronouns consist of less syntactic structure than strong pronouns. Strong pronouns are considered to possess three structural layers: a CP layer, a  $\Sigma$ P layer and an IP layer all indexed with the diacritic  $_N$  to indicate that these relate to the nominal domain and not the clausal domain. The head of  $C_NP$  is the locus of referential features and case; the head of  $\Sigma_NP$  contains polarity and focus features; and the head of  $I_NP$  contains agreement features, much like its clausal brethren. As can be seen from (6), weak pronouns and clitic pronouns are seen as reduced variants of strong pronouns.



Without entering the details of their analysis, the reduced structure in the deficient pronouns (both weak and clitic) are argued by C&S to have the following consequence:

Deficient pronouns:

- (i) **cannot be prosodically focused**
- (ii) **cannot occur in all of the regular argument positions**
- (iii) **cannot be coordinated**
- (iv) **cannot be modified**

Properties (i) and (ii) accord well with the behavior of Tagalog 2P clitics which cannot be prosodically focused, as shown in (7), and, as 2P elements, clearly cannot appear in regular argument positions. (We return to the inability of clitic pronouns to take narrow focus later in the chapter.)

- (7) \*Nag-lútò=[SILA]<sub>F</sub>  
AV.BEG-cook=3P.NOM  
(can only be interpreted as ‘[They cooked]<sub>F</sub>’)

Properties (iii) and (iv) require more discussion as the relevant data are more complex. In support of a general constraint against clitic conjunction we find evidence from Romance languages, where verb–adjacent clitics have neither the possibility for conjunction nor modification. Compare the French sentences in (8) and (9). In the (a) examples we find grammatical coordination and modification of full NPs and in the (b) examples we see ungrammatical instances of the same operations with verb-adjacent object proclitics. Kayne (1975) employed this data to argue that clitics must adjoin to V. Conjunction can then be ruled out by the fact that clitics cannot comprise syntactic constituents on their own.

- (8) a. Je      connais Jean et      Marie                      *French*  
1S.NOM know J.    CONJ M.  
‘I know Jean and Marie.’

- b. \*Je le et la=connais  
 1S.NOM 3S.M.ACC CONJ 3S.F.ACC=know
- (9) a. Il ne connaît que nous deux *French*  
 3S.NOM knows only 1P.ACC two  
 ‘He only knows us two.’
- b. \*Il nous=deux=connâit  
 3S.NOM 1P.ACC=two=knows

We can *a priori* distinguish two types of conjunction: conjunction of a pronominal with a full NP and coordination between two pronominals. Conjunction of a pronominal with a full NP appears to be allowed in Tagalog, as shown in (10)-(11).<sup>3</sup>

- (10) Kaya’t samá-han=niyo kami at aḡ=maḡa=hosts...  
 so join-LV=2P.GEN 1P.NOM CONJ NOM=PL=hosts  
 ‘So joins us and the hosts...’<sup>4</sup>
- (11) Kawáwá=naman kami at aḡ=maḡa=túlad=námin=ḡ marino  
 pitiful=EMPH 1P.NOM CONJ NOM=PL=similar=1P.GEN=LNK sailor  
 ‘We and are fellow sailors are truly pitiful...’<sup>5</sup>

However, the pronouns in such examples are ambiguous between 2P clitics and free pronouns. Recall that all nominative pronominal clitics except the second person singular have homophonous free counterparts. In order to test whether clitic pronouns can be conjoined we must look at either the second person singular nominative or the genitive clitic pronouns, whose free counterparts are phonologically distinct. In (12), we find an example of the first case with a genitive pronoun and in

<sup>3</sup> Zribi-Hertz & Mbolatianavolana (1999:177) claim that examples similar to those in (10)-(12) in Malagasy are the result of “augmentation” and on par with parentheticals. But their argument crucially rests on the lack of case on the second conjunct, and thus cannot be extended to the above data. Note also that Tagalog *at* is solely a conjunction and never a comitative preposition as in languages where conjunctions are formed from comitative phrases.

<sup>4</sup> From: <http://www.abs-cbn.com/entertainment/archives-perfectmoments.aspx>

<sup>5</sup> From: <http://www.ufs.ph/tinig/marapr01/03040128.html>

(13) we find an example of the unambiguous second person singular clitic *ka* conjoined with a full DP.

(12) Péro ngayon ay hindi na kami aṅ=na-kí~kinábang  
 but now TOP NEG LNK 1P.NOM NOM=AV.BEG-INCM~profit

sa=maṅa=p<in>ag-hiráp-an=**námin** at **naṅ=ámin=ṅ** maṅa=**ninúno**  
 OBL=PL=<BEG>TR-difficult-LV=1P.GEN CONJ GEN=1P.GEN=LNK PL=ancestor  
 ‘But now, we can no longer profit from that which our ancestors toiled for.’<sup>6</sup>

(13) *Nice pictures* péro bákit walà=**ka** at **si=Mystica?**  
 Nice pictures but why NEG.EXT=2S.NOM CONJ P.NOM=M.  
 ‘Nice pictures but why aren’t you and Mystica there?’<sup>7</sup>

Note however that in both of the above examples second position happens to coincide linearly with the post-predicate position of arguments. When a potential clitic host precedes the predicate, second position is linearly differentiated from argument position. We can call this position unambiguous 2P to differentiate it from the linearly ambiguous post-predicate position. For conjoined clitics in unambiguous 2P the data is somewhat more equivocal. Examples such as the following can be found but are far from commonplace. In (14)-(16), the conjoined clitics follow negation and in (17) and (18) they follow an adjunct interrogative. Although examples may be found relatively easily in casual written texts, speakers may hesitate in accepting them as fully grammatical.

(14) hindi=**ko** at **naṅ=maṅa=kasámahan=ko=ṅ** **Filipina** kailanman  
 NEG=1S.GEN CONJ GEN=PL=colleague=1S.GEN=LNK Filipina ever

s<in>írà-Ø aṅ=tiṅin naṅ=iba sa=ámin  
 <BEG>destroy-PV NOM=view GEN=other OBL=1P.GEN  
 ‘Neither I nor my Filipina colleagues ever destroyed the view of others towards us.’<sup>8</sup>

<sup>6</sup> From: [http://www.forestpeoples.org/documents/ifi\\_igo/wb\\_ips\\_philippines\\_may00\\_eng.shtml](http://www.forestpeoples.org/documents/ifi_igo/wb_ips_philippines_may00_eng.shtml)

<sup>7</sup> From: [www.a-pinoy-in-nz.blogspot.com/.../day-at-beach-and-seaworld.html](http://www.a-pinoy-in-nz.blogspot.com/.../day-at-beach-and-seaworld.html)

<sup>8</sup> From: [www.nursesthoughts.com/2007\\_09\\_01\\_archive.html](http://www.nursesthoughts.com/2007_09_01_archive.html)



- (15) Nápaka-sarap! péro alam=ko hindi=**ka**, at **si=Ishi**,  
 INTNS-delicious but know=1S.GEN NEG=2S.NOM CONJ P.NOM=Ishi  
 ma-sá~sarap-an don.  
 NVL-INCM-delicious-LV there  
 ‘So delicious! But I know that you and Ishi wouldn’t find it tasty.’<sup>9</sup>
- (16) hindi=**ka** at **sino=man díto** ta~tamá-an naṅ=maṅa=s<in>á~sábi=ko  
 NEG=2S.NOM CONJ who=ever here INCM~effect-LV GEN=PL=<BEG>INCM~say=1S.GEN  
 ‘What I say here won’t effect you and whoever else here.’<sup>10</sup>
- (17) Kelan=**ka** at **si=Jet**, **kasáma=na=rin si=Spidey**  
 when=2S.NOM CONJ P.NOM=J. with=CMP=also P.NOM=S.  
 ma-gá~gawì díto sa=dáko=ḡ Hilágà?  
 NVL.PV-INCM~direction here OBL=area=LNK north  
 ‘When are you and Jet along with Spidey going to come by the North?’<sup>11</sup>
- (18) kailan=**ka** at **maṅa=ka-sáma=mo=ḡ**  
 when=2S.NOM CONJ PL=co-accompany=2S.GEN=LNK  
**maṅa=twisted na sina=fromtoronto at dominus\_iesus** Ø-ti~tinò?  
 PL=twisted LNK P.NOM.PL=f. CONJ d. AV-INCM~sane  
 ‘When will you and your twisted friends, *fromtoronto* and *dominus iesus*,  
 become sane?’<sup>12</sup>

Conjunction of two pronominal clitics is judged even worse by speakers yet rare examples of this can also be found, as in (19), an example from a transcribed oral interview, and (20). The conjoined genitive pronouns in (19) follow negation in the pre-predicate field and are in turn followed by a nominative clitic within the clitic cluster, thereby eliminating any doubt that all of these pronouns are bona-fide 2P clitics.

- (19) akálà=ko hindi=na=ako maká~ka-balik péro  
 thought=1S.GEN NEG=ALRD=1S.NOM AV.NVL~INCM-return but

<sup>9</sup> From: <http://yduj.multiply.com/>

<sup>10</sup> From: [www.timog.com/forum/showthread.php?t=19245&page=6](http://www.timog.com/forum/showthread.php?t=19245&page=6)

<sup>11</sup> From: <http://kwentongtambay.com/?p=761>

<sup>12</sup> From: <http://www.gov.ph/forum/thread.asp?rootID=76196&catID=24&page=10>

hindi=**nila at ninyo**=ako p<in>a-bayá-an.  
 NEG=3P.GEN CONJ 2P.GEN=1S.NOM <BEG>CAU-neglect-LV  
 ‘I thought I wouldn’t be able to return anymore but you and they  
 didn’t neglect me.’<sup>13</sup>

- (20) kíta=naman=táyo pag free=**ka at ako at maḡa**=superfriends  
 see=SWTCH=1P.NOM when free=2S.NOM CONJ 1S.NOM CONJ PL=superfriends  
 ‘Let’s see each other when you and I and the superfriends are free.’<sup>14</sup>

For most speakers, however, such clitic conjunction is ungrammatical, as reflected by the judgments shown in (21)a. For these speakers, to express the intension of (21)a, it is necessary to employ a single pronoun which matches the sum of the person features. Thus, the sum of the features of the first person exclusive plural and the second person plural is the first person inclusive plural, as given in (21)b.

- (21) a. \*hindi=kami at kayo manánálo  
 NEG=1P.NOM CONJ 2P.NOM AV.INCM:win  
 (For, ‘We and you won’t win.’)  
 b. hindi=táyo manánálo  
 NEG=1+2P.NOM AV.INCM:win  
 ‘We won’t win.’

Note, however, that for no speaker does this constraint apply to free pronominals (which are generally otherwise homophonous with their clitic counterparts), as shown in (22), where the pronouns are in predicate position in a cleft-like construction with the verb embedded within a nominative phrase.

- (22) kami at kayo aḡ=manánálo  
 1P.NOM CONJ 2P.NOM NOM=AV.INCM:win  
 ‘We and you are the ones who will win.’

<sup>13</sup> From: <http://www.pep.ph/articles/17174/Ogie-Alcasid-endured-cramps-from-exhaustion-to-finish-his-concert>

<sup>14</sup> From: [www.tinakuting.multiply.com](http://www.tinakuting.multiply.com)

Having seen full NPs being “coerced” into clitic position by conjunction in (14)-(18) (for those speakers who allow it), we may ask whether or not clitics may also be “coerced” into argument position by the same process. Sityar (1989) claims that apparent clitic pronouns can appear in non-clitic positions, offering the position of the nominative clitic in example in (23) as evidence. As discussed by Billings (2005), the grammaticality of such examples is highly questionable as they appear to be rejected by many speakers. It should be further added here that the nominative clitic in Sityar’s example is positioned ambiguously; it can be interpreted as being in either the position of a nominative argument or in a “delayed” clitic position. If such pronouns were indeed treated as full NP arguments then we expect them to also be able to take the canonical position of the nominative argument, following the genitive agent phrase. The ungrammatical sentence in (24) makes clear that, inasmuch as (23) can be judged grammatical by some speakers, the nominative pronoun should be considered to be in delayed clitic position rather than a true argument position.

(23) ?Hindi=ko na=kítà-Ø **siya** kahápon  
 NEG=1S.GEN NVL.BEG-see-PV 3S.NOM yesterday  
 ‘I didn’t see her yesterday.’ (Sityar 1989:16)

(24) \*Hindi na=kítà-Ø naŋ=pulis **siya** kahápon  
 NEG NVL.BEG-see-PV GEN=police 3S.NOM yesterday  
 (For, ‘The police didn’t see her yesterday.’)

Nonetheless, conjunction of pronouns does appear to allow for exceptional placement in argument position, as seen in (25) (although this sentence is judged as less than fully grammatical by speakers, as indicated). Here the nominative pronoun appears after the genitive phrase, in the normal position of a nominative NP. As shown by the clearly ungrammatical variant in (26), such positioning is impossible for an unconjoined pronoun, which must follow the first available host in its domain.

- (25) <sup>?</sup><in>alipustà-Ø naḡ=maḡa=Egipcio **kami** **at** **aḡ=ámiḡ=maḡa=magúlaḡ**  
 <BEG>afflict-PV GEN=PL=Egyptian 1P.NOM CONJ NOM=1P.GEN:LNK=PL=parent  
 ‘The Egyptians afflicted us and our fathers.’ (Numbers 20:15)
- (26) <in>alipustà-Ø[=kami] naḡ=maḡa=Egipcio [**\*kami**]  
 <BEG>afflict-PV=1P.NOM GEN=PL=Egyptian  
 ‘The Egyptians afflicted us.’

It is clear that in this case the conjoined pronoun is free and not clitic. This can again be tested by looking at the second person singular and the genitive set. In (27), we see that when the pronoun in (25) is replaced with the second person singular, the free form is marginally possible, as in (27)a and the clitic form is totally ungrammatical, as shown by (27)b. The example in (28) demonstrates the same point with an unambiguously clitic genitive pronoun. Again conjunction of the clitic in argument position is ungrammatical.

- (27) a. <sup>?</sup><in>ulúḡ-an naḡ=maḡa=Egipcio **ikaw** **at** **aḡ=pínsan=mo**  
 <BEG>help-LV GEN=PL=Egyptian 2S.NOM CONJ NOM=cousin=2S.GEN  
 ‘The Egyptians helped you and your cousin.’
- b. **\*t**<in>ulúḡ-an naḡ=maḡa=Egipcio=**ka** **at** **aḡ=pínsan=mo**  
 <BEG>help-LV GEN=PL=Egyptian=2S.NOM CONJ NOM=cousin=2S.GEN
- (28) **\*t**<in>ulúḡ-an aḡ=pínsan=mo=**nila** **at** **naḡ=maḡa=rebélde**  
 <BEG>help-LV NOM=cousin=3P.GEN CONJ GEN=PL=rebel  
 (For, ‘They and the rebels helped your cousin.’)

The facts discussed above are summarized in Table 2.1 according to construction.

**Table 2.1.** Summary of coordination facts

Configuration	Acceptability
clitic & DP in ambiguous 2P/arg position	✓
clitic & DP in unambiguous 2P	?
clitic & DP in unambiguous arg position	*
clitic & clitic	*
free pronoun & DP in unambiguous 2P	*
free pronoun & DP in unambiguous arg position	?

Although the coordination judgments are delicate and require further testing, the picture which emerges from the available data hints at interesting interactions between clitics and full DP arguments. While it seems fair to say that speakers generally prefer “inclusive modification” (to be discussed directly below) over pronominal coordination, a coordinated clitic appears to be fully acceptable when its position is ambiguous between (linear) argument position and 2P. Despite the attestations in (14)-(18), coordination of 2P clitics in unambiguous clitic position is regularly judged as degraded in careful speech and completely ungrammatical in argument position. Likewise, coordination of free pronouns in unambiguous argument position is also judged to be less than fully acceptable, despite rare attestations. Intriguingly, the acceptability of coordinated clitics in ambiguous positions suggests that the constraints on 2P positioning as well as those on the contiguity of conjuncts are evaluated on the surface. In other words, clitic coordination is licensed so long as the clitic can maintain its canonical surface positioning in 2P and the argument can remain in-situ. The significance of this will be examined more carefully in §2.4.1.

This brings us to a functionally similar construction which can be termed “inclusive modification” (cf. S&O 1972:116). Inclusive modification is commonly employed to express a proper subset relation between two referents in Tagalog. It involves a pronominal followed by a genitive phrase with the latter’s reference being included in the pronouns person/number features. The genitive phrase is typically a full DP and in older Tagalog could also be a pronoun. Because inclusive modification with two pronouns is not attested in present-day Tagalog we will restrict our examples to those with a full DP genitive phrase. Note that the inclusive phrase is always assigned genitive case, even when it is interpreted as part of a nominative phrase, as in (29). Note also that the pronoun must include the features of the inclusive phrase, as demonstrated by the ungrammatical (30). Furthermore, the inclusive phrase can only

be licensed by a pronoun and not by a full DP, as shown by (31). To express the intention of (31) it is necessary to employ a prepositional phrase, as in (32). In the free translation of the examples below the inclusive phrase is introduced by the proper subset symbol  $\supset$ .

- (29) d<um>atiŋ=kami            ni=Pablo  
 <AV.BEG>arrive=1P.NOM P.GEN=Pablo  
 ‘We  $\supset$  Pablo arrived.’
- (30) \*d<um>atiŋ=ako            ni=Pablo  
 <AV.BEG>arrive=1S.NOM P.GEN=Pablo
- (31) \*d<um>atiŋ    aŋ=lahat ni=Pablo  
 <AV.BEG>arrive NOM=all P.GEN=Pablo  
 (For, ‘Everyone  $\supset$  Pablo arrived.’)
- (32) d<um>atiŋ    aŋ=lahat, pati        si=Pablo  
 <AV.BEG>arrive NOM=all including P.NOM=Pablo  
 ‘Everyone arrived, including Pablo.’

Although not directly relevant here, there is a subtle but noticeable semantic difference between clitic & DP coordination and the inclusive construction. The referents of a coordinated phrase may be fully individuated while those in the inclusive construction often receive a group interpretation. In the minimal pair given in (33), the (a) sentence is more appropriate for a looser connection between the hearer and Mikki while the (b) sentence is more appropriate in addressing a married couple, or some such other collective entity.

- (33) a. Kamusta=ka at si=Mikki?            b. Kamusta=kayo ni=Mikki?  
 how=2S.NOM CONJ P.NOM=M.            how=2P.NOM P.GEN=M.  
 ‘How are you and Mikki?’<sup>15</sup>            ‘How are you  $\supset$  Mikki?’

<sup>15</sup> From: [www.mahriz.multiply.com](http://www.mahriz.multiply.com)

As was also seen above with the coordination data, DPs in the inclusive construction can also occasionally be found in “coerced” clitic positions. We see this in (34) and (35) where the clitic and the following inclusive phrase are both in unambiguous 2P, following an interrogative in the pre-predicate field.

- (34) *Wala=ŋ            may naká~ka-alam*  
 NEG.EXT=LNK EXT AV.BEG~INCM-know  
*kuŋ=saan=sila            naŋ=pamilya=niya p<um>unta*  
 COMP=where=3P.NOM GEN=family=3S.GEN <AV.BEG>go  
 ‘There was no one who knew where they ⊃ his family went.’<sup>16</sup>
- (35) *isip na isip=siya            ŋayon kuŋ paano=nila naŋ=kanya=ŋ            asáwa*  
 think LNK think=3S.NOM now COMP how=3P.GEN GEN=3S.GEN=LNK spouse  
*gá~gaw-iŋ            ma-saya aŋ=birthday naŋ=kanila=ŋ            anak*  
 INCM~do-PV:LNK ADJ-happy NOM=birthday GEN=3P.GEN=LNK child  
 ‘Now he thought and thought about how they ⊃ his spouse would make their child’s birthday happy.’

Unlike in coordination, the associated pronoun may be disassociated from the inclusive phrase, which can be placed in argument position, as shown in (36).

- (36) *hindi=sila            d<um>atiŋ            naŋ=mom=ko*  
 NEG=3P.NOM <AV.BEG>arrive GEN=mom=1S.GEN  
 ‘They ⊃ my mom didn’t arrive.’<sup>17</sup>

In voices other than the actor voice where there is an expected genitive marked agent in the post-predicate field, inclusive modification is somewhat marked, perhaps because of the ambiguity which it entails. Speaker judgments for inclusive modification with a patient voice predicate are given in (37). In (37)a, the favored of the three variants, the pronoun and the associated inclusive phrase all appear in 2P. In (37)b, which is slightly degraded, the pronoun is in clitic position and the inclusive

<sup>16</sup> From: [www.jhayb08.wordpress.com](http://www.jhayb08.wordpress.com)

<sup>17</sup> From: <http://gergermae.multiply.com/journal>

phrase appears in argument position. Placing both the pronoun and the inclusive phrase in argument position, as in (37)c, is completely ungrammatical. Thus, what is common to the grammatical variants is that the pronoun must appear in 2P, preferably forming a constituent with the inclusive phrase.

*Inclusive modification – 2P*

- (37) a. na-kítà-Ø=**kami**      **naŋ=maŋa=kaibigan=ko**    naŋ=pulis  
 NVL.BEG-see-PV=1P.NOM    GEN=PL=friend=1S.GEN    GEN=police  
 ‘The police saw me and my friends.’

*Inclusive modification – disassociated 2P clitic*

- b. ?na-kítà-Ø=**kami**      naŋ=pulis    **naŋ=maŋa=kaibigan=ko**  
 NVL.BEG-see-PV=1P.NOM    GEN=police    GEN=PL=friend=1S.GEN

*Inclusive modification – argument position*

- c. \*na-kítà-Ø      naŋ=pulis    **kami**    **naŋ=maŋa=kaibigan=ko**  
 NVL.BEG-see-PV    GEN=police    1P.NOM    GEN=PL=friend=1S.GEN

As regards direct modification of pronouns, we find both numeral and other lexical modifiers linked to clitic pronominals. As shown in (38) and (39), modification of clitics by numerals is permitted. In (38), the pronoun is in unambiguous 2P as it follows the interrogative and precedes the predicate. Similarly, in (39), the pronoun follows a pre-predicate modifier in a position where full NP arguments are not permissible.

- (38) Saan=**kayo=ŋ**      **lima** nag-tuloy      matápos    kayo=ŋ  
 where=2P.NOM=LNK    five    AV.BEG-continue after    2P.NOM=LNK  
 maka-babà      naŋ=Bus?  
 AV.NVL-descend    GEN=bus  
 ‘Where did you five continue to after getting off the bus?’<sup>18</sup>

<sup>18</sup> From: [http://www.lawphil.net/judjuris/juri1980/dec1980/gr\\_26944\\_45\\_1980.html](http://www.lawphil.net/judjuris/juri1980/dec1980/gr_26944_45_1980.html)



- (39) **sabay=ninyo=ŋ dalawa=ito=ŋ linís-in**  
 simultaneous=2P.GEN=LNK two=this=LNK clean-PV  
 ‘the both of you clean this at the same time.’<sup>19</sup>

The numeral can also appear disassociated from the clitic and connected to the predicate by the linker, as in (40) and (41). Note that, unlike the case of inclusive modification, the linker indicates that the disassociated material in the post-predicate field is not in argument position but rather has a direct dependency on the predicate head. Similar modification of NPs by linking to the predicate is discussed by Kroeger (1993), Schachter (1996) and Kaufman (2009) under the rubric of floating quantification and secondary predication.

- (40) **Gusto=ko=laŋ=sila=ŋ ma-ka-harap-∅ na dalawa**  
 want=1S.GEN=only=3P.NOM=LNK NVL-CO-face-PV LNK two  
 ‘I just want to face them two.’<sup>20</sup>

- (41) **Ma-tagal=na=nila=ŋ t<in>angap-∅ na dalawa na hiwalay=sila**  
 ADJ-long.time=ALRD=3P.GEN=LNK <BEG>accept-PV LNK two COMP separate=3P.NOM  
 ‘The two of them have already long accepted that they are separated.’<sup>21</sup>

Examples of modification of 2P clitics by (non-numeral) lexical material are given in (42)-(45). In all cases, the modified pronouns follow an interrogative element or negation and precede the predicate, a position disallowed for full DP arguments.<sup>22</sup>

<sup>19</sup> From: [www.chp.dhs.lacounty.gov/pdf/ly\\_ta.pdf](http://www.chp.dhs.lacounty.gov/pdf/ly_ta.pdf)

<sup>20</sup> From: <http://www.gmanews.tv/story/129119/PEP-MTRCB-chief-glad-alls-well-between-Joey-Willie>

<sup>21</sup> From: <http://www.newsflash.org/2004/02/sb/sb003873.htm>

<sup>22</sup> The fact that genitive pronouns can be modified can give rise to ambiguity as a modifier phrase can be interpreted as modifying either the pronoun or the head noun from a single linear position. For instance, both the bracketings shown in (i) and (ii) are available for the same string.

(i) aŋ=maŋa=[kaibígan[=námiŋ wala=ŋ péra]]  
 NOM=PL=friend=1P.GEN:LNK NEG.EXT=LNK money  
 ‘The friends of [we who have no money].’

(ii) aŋ=maŋa=[[kaibígan=námiŋ] wala=ŋ péra]  
 NOM=PL=friend=1P.GEN:LNK NEG.EXT=LNK money  
 ‘Our friends who have no money.’

- (42) Kailan=**kayo=ŋ**    **maŋa=dalawa=ŋ** **badiŋ**    mag-pá~pa-kasal    kayà?  
 when=2P.NOM=LNK PL=two=LNK    homosexual AV-CAU~INCM-marry SPEC  
 ‘When will you two homosexuals get married?’<sup>23</sup>
- (43) tiŋn-an=nátin    kuŋ haŋgang saan=**kayo=ŋ**    **maŋa=INC** maká~ka-ratiŋ  
 look-LV=1+2P.GEN COMP until    where=2P.NOM=LNK PL=INC    AV.NVL~INCM-arrive  
 ‘Let’s see how far you INC (*Iglesia ni Cristo*) get.’<sup>24</sup>
- (44) Hindi Ø-á~alis    ‘yon naŋ=báhay naŋ=hindi=**kami=ŋ**    **maŋa=bátà**  
 NEG    AV-INCM~leave that    GEN=house GEN=NEG=1P.NOM=LNK PL=child  
 na-abút-an    naŋ=péra.  
 NVL.BEG-hand.over-LV GEN=money  
 ‘He wouldn’t leave the house without handing us kids some money.’<sup>25</sup>
- (45) hindi=**nátin**    **maŋa=banyágà** na-rá~ramdam-an    aŋ=maŋa=pátáy-an  
 NEG=1+2P.GEN:LNK PL=foreigner    NVL.BEG-INCM~feel-LV NOM=PL=kill-LV  
 ‘We foreigners don’t feel the killings.’<sup>26</sup>

Modification of pronominals is only possible with plural pronouns and is often accompanied by the plural marker *maŋa* as in the above examples, although this is not strictly necessary. As seen by (42), the modifier phrase is not required to be particularly light (cf. Billings 2005) and can contain both numerals and substantives. On the other hand, the modifier must be an NP and not a DP, as was the case with inclusive modification and coordination. I take this difference to be a critical factor in the apparent full acceptability of numerically modified pronouns in 2P, seen in (38) and (39), and NP modified pronouns in 2P, seen in (42)-(45), versus the reduced acceptability of pronouns coordinated with DPs in 2P as summarized in Table 2.1. This does not, however, account for the apparent full acceptability or near full

<sup>23</sup> From: <http://forum.gov.ph/thread.asp?rootID=91024&catID=6&page=14>

<sup>24</sup> From: [www.thebereans.net/forum2/showthread.php?t=38194&page=4](http://www.thebereans.net/forum2/showthread.php?t=38194&page=4)

<sup>25</sup> From: <http://www.palancaawards.com.ph/2006Batang%20Tundo%20by%20Rosario%20Torres-Yu%201st%20prize.html>

<sup>26</sup> From: <http://74.125.45.132/search?q=cache:er9AxZ-dDB8J:www.siampinoy.com/forum2/index.php+http://www.siampinoy.com/forum2/index.php/topic,565.msg28961.html&hl=en&ct=clnk&cd=1&gl=us&client=safari>

acceptability of inclusive DPs in unambiguous 2P (§2.4.1). Finally we note that, unlike inclusive modification, linking modification appears to license positioning in argument position, as shown in (46)a, where the modified pronominal is in unambiguous argument position. This position is otherwise unavailable to a bare pronoun, which must follow the adverbial directly, as shown by the judgments for the two pronoun positions in (46)b.<sup>27</sup>

- (46) a. Ma-dalas    na-bá~bansag-aŋ            ma-árte    **kami=ŋ**    **maŋa=Atenísta**  
 ADJ-frequent    BEG-INCM~name-LV:LNK    ADJ-snob    1P.NOM=LNK    PL=Atenista  
 ‘We Atenistas are frequently called snobby.’<sup>28</sup>
- b. Ma-dalas[=kami=ŋ]            na-bá~bansag-aŋ            ma-árte    [\*kami]  
 ADJ-frequent=1P.NOM=LNK    BEG-INCM~name-LV:LNK    ADJ-snob    1P.NOM  
 ‘We are frequently called snobby.’

In sum, the data presented above from both coordination and modification presents problems for the C&S typology. Tagalog pronouns in 2P clearly fall into the clitic category of their taxonomy as given in (5) above, as these pronouns cannot appear in argument position. At the same time, these 2P clitics allow modification and to a certain extent, coordination which C&S predict to be illicit not only for clitics but for the larger category of deficient pronouns. According to their taxonomy clitics are further restricted by the following constraints:

Clitic pronouns:

- (i) **cannot bear stress**
- (ii) **cannot occupy any argument positions**

Although C&S do not propose principled links between morphosyntactic characteristics and phonological behavior, they do argue that pronouns which cannot

---

<sup>27</sup> Unfortunately, we cannot know if the pronominals coerced to argument position are clitics or free pronouns because modification is only felicitous with plural pronouns and the clitic free pronoun distinction is only formally diagnosable by the second person singular. On analogy with the coordination data, however, it seems safe to assume that these are free pronouns and not clitics.

<sup>28</sup> From: [www.lastrenaissance.livejournal.com/30267.html](http://www.lastrenaissance.livejournal.com/30267.html)

occupy argument positions cannot bear stress. Again contrary to their typology, Tagalog 2P clitics can bear stress and can even bear intonation phrase prominence (albeit not prosodic focus).

All the distributional facts reviewed above are summarized in Table 2.2 below. The leftmost column indicates the general type of construction and the grammaticality judgments reported. The distribution of the pronominal and its associate is given schematically in the righthand column. The (a) items represent both clitic and associate in unambiguous 2P; the (b) items in unambiguous argument position; the (c) items represent a discontinuous ordering with the clitic in 2P and its associate in argument position; the (d) items represent a surface position of clitic and DP which is linearly ambiguous between argument position and 2P and the (e) items represent the possibility of a free pronominal in argument position with its associate.

**Table 2.2.** Summary of data on coordination, inclusive and linking modification

Construction	2P	Arg. position
(i) Coordination		
a. ?	=cl & DP	...
b. *		... =cl & DP
c. *	=cl	... & DP
d. ✓	=cl	& DP
e. ?		... pron. & DP
(ii) Inclusive		
a. ✓	=cl $\supset$ DP	...
b. *		... =cl $\supset$ DP
c. ✓	=cl	... $\supset$ DP
d. ✓	=cl	$\supset$ DP
e. *		pron. $\supset$ DP
(iii) Modification		
a. ✓	=cl LNK DP	...
b. (see fn.27)		... =cl LNK DP
c. * <sup>29</sup>	=cl	... LNK DP
d. ✓	=cl	LNK DP
e. ✓		... pron. LNK DP

<sup>29</sup> Recall that discontinuity is licensed here but only when the modifier is linked to the predicate head rather than in argument position.

In the next section we examine the theory of clitics put forth by Condoravdi & Kiparsky (2001) which predicts different correlations between morphosyntactic and phonological properties and again evaluate this theory against the Tagalog data.

### 2.2.3 *Kiparsky & Condoravdi (2001, 2004)*

Condoravdi & Kiparsky (2001) (henceforth C&K) argue for a three-way distinction in the morphology of person marking on the basis of Greek dialect data. Unlike the proposal of C&S reviewed above, C&K seek to tie the phonological behavior of clitics with their positional patterns. C&K posit three basic categories representing different places of attachment for person markers, making their theory more nuanced than that of analyses such as Halpern & Fontana (1994) which only admit two types of clitics. At the same time, they still maintain the claim that clitics can be classified cross-linguistically into discrete categories, making their theory more restrictive than that of Janse (1998), who posits a cline of morphological attachment without discrete categories. Importantly for our purposes, C&K take into account the position and X-bar status of the clitic host in addition to clitic behavior with coordination of the host. C&K's three types are shown in Table 2.3, alongside their attendant characteristics.<sup>30</sup> Note that the third category is not very clitic-like and rather describes the outermost level of affixation, considered to take place by C&K within the lexicon.

---

<sup>30</sup> C&K do not give examples of X<sup>0</sup> clitics which change polarity, proclitic when a preceding host is available and enclitic otherwise (i.e., Tobler-Mussafia clitics). But this possibility may be adduced in their system on fairly standard assumptions. Switching polarity for affixes, on the other hand, is assumed to be impossible because of the lexical nature of prefixing versus suffixing, although rare examples of such affixes do exist in the literature (see Noyer 1993 for Huave and Fulmer 1997 for Afar) and certain theories make provisions for this (e.g., Miller 1992, Embick & Noyer 1999).

**Table 2.3.** Condoravdi & Kiparsky's (2001) clitic typology

	<b>X<sup>max</sup> clitics</b>	<b>X<sup>0</sup> clitics</b>	<b>Word-level affixes</b>
Area of grammar	syntax	syntax	Lexicon
Locus of attachment	TP/CP	[±FIN] V <sup>0</sup>	[+FIN] V
LexCat selection	×	✓	✓
Sharing with CONJ	✓	✓	×
Lexical phonology	post-lex	post-lex word	word phonology

X<sup>max</sup> clitics need not be 2P clitics, although presumably, all 2P clitics must be X<sup>max</sup> clitics in such a tripartite analysis. For C&K, 2P effects are derived from attachment of a clitic to a higher functional phrase (e.g., TP, CP, etc.) and canonical XP movement of the clitic host into [Spec,TP] or [Spec,CP] (or merger of a complementizer host in C). If these processes do not yield a host for the clitic then a process of Prosodic Inversion (to be discussed further below) takes place to provide a suitable host.

In the dialectal and historical Greek data to which C&K restrict themselves, Type A dialects (Eastern dialects) possess pronominal X<sup>max</sup> clitics which are analyzed as attaching to TP. Consequently, they follow complementizers, negation, modal particles, foci and wh- phrases when they are clause initial. When none of the above elements are available as hosts, clitics follow the verb as a result of Prosodic Inversion. These clitics participate in post-lexical, but not word-internal, phonology and can be shared by conjoined hosts (i.e., as in the structure [X & Y]=*cl*).

In Type B dialects (Pontic, Kozani) clitics are head-adjoined to V, rather than to a functional projection. They can be either enclitics (as in Pontic) or proclitics (as in Kozani). Evidence that these clitics are neither adjoined to T nor word-level affixes comes from the fact that they attach to an infinitive verb in the presence of a finite auxiliary. They differ from word affixes in allowing conjunction of a verbal host under a single clitic, as with X<sup>max</sup> clitics. Finally, C&K show phonological evidence that these clitics are not part of the lexical word but rather the post-lexical word.

In Type C dialects (Western dialects, including Standard Modern Greek) the person markers in question are analyzed as word-level affixes, meaning that they attach to words to form larger words.<sup>31</sup> They attach lexically to the left of a finite verb in unmarked declaratives and cannot be shared by conjoined verbs (i.e., \*[V & V]=cl). Their restriction to finite verbs, and therefore similarity to agreement morphology, is the basis of their treatment as lexical and not syntactic items. They furthermore participate in word phonology. Lexical affixes are presumably specified as prefixes and suffixes in their lexical entry and cannot change their polarity (Left/Right edge alignment) based on the prosodic environment external to the word. Their polarity can however depend on the mood of the sentence, which, under several analyses, may trigger verb movement around the clitic. For C&K, the main distinction between  $X^0$  and word-level affixes is that the former attach in the syntax while the latter attach in the lexicon, a distinction not made by previous theories.

For C&K, morphological attachment is tied to prosodic attachment by virtue of the separation of lexical and post-lexical phonology into two different strata and the presence of different levels of morphophonology within the word itself (Kiparsky 2003). This seems to predict that 2P elements, as  $X^{\max}$  clitics, should not be able to participate in lexical phonology. However, many cases exist of 2P clitics partaking in what appear to be low-level phonological rules. Halpern (1995 chap. 5) attempts to demonstrate that the clitics in these cases are not true 2P clitics but rather examples of “extended inflection” which attaches to the head of an initial constituent. It is far from clear that all such cases can be argued away on similar grounds. In Ledo (Pamona-Kaili, Central Sulawesi, Austronesian), primary word stress falls on the penultimate syllable of the word without noticeable secondary stress. This can be demonstrated most clearly with the possessive enclitics, which may be disyllabic. In (47)a-c, we see

---

<sup>31</sup> Joseph (1988) also argues that Modern Greek person markers are affixes and not clitics.

a bare noun and the effect on stress when a monosyllabic and disyllabic genitive enclitics are added. Note that, crucially, there is no noticeable stress at all on (47)c, showing that the enclitic must be considered part of the minimal stress domain.

- (47) a. *banúa*                      b. *banuá=ku*                      c. *banua=kámi*                      *Ledo*  
house                              house=1S.GEN                      house=1+3.GEN  
‘house’                              ‘my house’                              ‘our house’ (Esser 1934:5)

The syntax of possessive enclitics is consistent with them being either lexical affixes or  $X^0$  clitics in C&K’s framework. Problematically, though, the 2P aspectual clitic *=mo* has the same effect on stress. The positioning domain of the aspectual clitic excludes the preverbal subject but includes CP elements such as interrogatives.<sup>32</sup> This is shown in the comparison between (48)a and b, where the stress is shifted to the final syllable of the verb with the addition of the clitic.

- (48) a. *Ía ne-gúru*                      b. *Ía ne-gurú=mo*                      *Ledo*  
3S AV.BEG-study                      3S AV.BEG-study=ALRD  
‘He’s studying.’                      ‘He’s already studying.’

That *=mo* is truly a 2P clitic can be seen by its positioning after the interrogative in (49) (attachment to the verb is ungrammatical here). Just as with verbal attachment, when the clitic attaches to the interrogative, stress is again shifted to the right.<sup>33</sup>

---

<sup>32</sup> A similar situation obtains with V-adjacent clitics in Greek dialects (see C&K 2002). This can be handled trivially by assuming that subjects are adjoined topics, a move which is independently justified in *Ledo*, which has apparently only recently begun to treat topicalized pre-verbal subjects as pragmatically unmarked.

<sup>33</sup> Typically, stress assignment is thought to apply post-lexically to produce secondary stress or to resolve clashes which arise between the word-level and higher levels. The *Ledo* facts suggest that post-lexical stress must be able to assign the sole primary stress within a word as there is no clear secondary stress in the language. Although this is possible in principle it is an unusual state of affairs and potentially problematic when taking a wider range of facts into account.



- (49) Nakuyá=**mo** lédo ne-gúru ía? *Ledo*  
 why=ALRD NEG AV.BEG-study 3S  
 ‘Why is he not studying anymore?’

This forces an analysis in which the primary stress associated with each phonological word is derived post-syntactically after attachment of  $X^{\max}$  clitics. Although strictly speaking not impossible, this disallows the bracket erasure which generally typifies post-lexical processes.

Furthermore, Esser (Esser 1934:4) reports an optional shortening process which is triggered particularly by 2P clitics, shown in (50). Such a process would again be surprising if =*mo* must be parsed post-lexically as an  $X^{\max}$  clitic since this shortening does not appear to operate more generally in the language. Again, treating this 2P morpheme as an  $X^{\max}$  clitic forces the post-lexical phonology to peer more deeply than expected into word structure.

- (50) /ne-tuwu=**mo**/ ⇒ [netú:mo] *Ledo*  
 AV.BEG-live=ALREADY

We also find difficulties in applying C&K’s theory of discrete clitic types to 2P elements on morphosyntactic grounds. Kaufman (to appear) demonstrates the 2P status of absolutive clitics throughout the South Sulawesi family (see also Friberg 1988 and Basri 1999) showing their syntactic behavior in a broad range of environments. In (51), from Mandar, we see that the absolutive clitic is not head adjacent but must instead follow the entire NP predicate. In (52), from Selayar, we see that the clitic must follow a locative phrase when fronted to a focus position as in (52)b. Finally, in (53), again from Mandar, we see that the clitic can be separated from the verb by any number of intervening elements when the left periphery is occupied by multiple potential hosts. These basic generalizations hold throughout the

South Sulawesi family and again force an analysis of these pronominals as X<sup>max</sup> clitics under C&K's theory.

- (51) Tongguru[\*=**a**'] ma-cowa[=**a**'] *Mandar*  
 teacher[=1s.A] STA-good  
 'I'm a good teacher. (Kaufman to appear: ex.132)
- (52) a. Tinro=**ko** ri=kadera b. Ri=kadera=**ko** tinro *Selayar*  
 sleep=2.A OBL=chair OBL=chair=2.A sleep  
 'You slept on a chair.' 'You slept on a chair.' (Basri 1999:250, ex.16a,b)
- (53) Andiaŋ=**i** pura melo' lamba sumobal i=Kaco' *Mandar*  
 NEG=3.A already want go sail PM=Kaco  
 'Kaco never wanted to go sail.' (Badudu 1990)

Interestingly, in all languages of this subgroup there exists a constraint against conjoined predicates sharing a clitic. This is exemplified by four languages in (54)-(57), where omission of either absolutive clitic results in ungrammaticality.<sup>34</sup> The impossibility of sharing a phrasal clitic cannot be explained easily within a framework which ties such syntactic properties of clitics directly to their positions in phrase structure as in C&K's approach (among others).<sup>35</sup>

- (54) ŋ-anre\* [= **a**] na ŋ-inuŋ=**a** *Selayar*  
 AV-eat=1s.A CONJ AV-drink=1s.A  
 'I ate and drank.' (Kaufman to appear:ex. 142)
- (55) K<um>ande\* [= **na**'] sola niso=to=**na**' *Duri*  
 <AV>eat=1s.A CONJ drink-also=1s.A  
 'I ate and drank.' (ibid. ex.143)
- (56) Ha=manne\* [= **ka**'] paŋ ha=menuŋ=**ka**' *Seko Padang*  
 NEG=eat=1s.A CONJ NEG=drink=1s.A  
 'I didn't eat or drink.' (ibid. ex.144)

<sup>34</sup> From (54)-(57), it may appear that the sequential reading of the conjunctions may be part of the problem here but this is in fact irrelevant to the general ban on clitic sharing in these languages.

<sup>35</sup> A rigid formalization of this type can be seen, among other places, in Miller's Criterion 1: "An item which cannot have wide scope over a coordination of hosts cannot be a postlexical clitic, and must be an affix" (Miller 1992:155).

- (57) Map-poloŋ\* [=ŋa'] anna mat-tunu=a' ayu *Mandar*  
 AV-chop=1s.A CONJ AV-burn=1s.A wood  
 'I chop and burn wood.' (ibid. ex. 145)

To sum up the preceding, we have seen evidence that 2P clitics are more closely attached to their hosts than allowed within C&K's theory given that these elements must be analyzed as  $X^{\max}$  clitics. 2P clitics can both effect low level phonological processes (as with primary stress placement and root truncation in Ledo) and require repetition under verb conjunction (as in the South Sulawesi languages). T

The true syntactic and prosodic correlates of clitic host conjoinability under a single clitic are still unclear. Looking beyond the obvious descriptive generalization that the inability of clitic sharing between conjoined hosts implies higher integration into the word, it is doubtful whether any universals can be drawn from a more detailed analysis. The variation found in Romance languages shows that V-adjacent clitics may either disallow clitic sharing, as in French (58) (cf. Miller 1992b), or allow it, as in European Portuguese (59).<sup>36</sup>

- (58) a. Je le=connais et je l'=aime bien *French*  
 1S.NOM 3S.M.ACC=know CONJ 1S.NOM 3S.M.ACC=like well  
 'I know him and I like him well.' (Wegmuller 1993:19; Reimsdijk 1995)
- b. \*Je le={connais et aime} bien  
 1S.NOM 3S.M.ACC=know CONJ like well (ibid.)
- (59) O Carlos disse que te={traz as segundas e leva as sextas} *EP*  
 the C. said that you=he.brings on Mondays and he.takes on Fridays  
 'Carlos said that he will bring you on Mondays and take you on Fridays.'

This type of difference has been analyzed as one between lexical affixes versus true cliticization (cf. Miller & Sag 1997). Note though that the lack of universal

<sup>36</sup> Rizzi (1986) analyzed the possibility of pronoun sharing as a diagnostic for (a unitary) clitic hood but this is clearly oversimplified in light of the more graduated picture obtained from data such as that from the Greek dialects.

agreement in affixal versus clitic properties makes this difficult to evaluate. More importantly, the same variation exists with 2P clitics. As seen earlier, in the South Sulawesi languages, sharing conjoined predicates under a single 2P clitic is impossible, as seen from (60) (repeated from (54)). Similar 2P clitics in Tagalog and other Philippine languages, on the other hand, freely allow conjunction under 2P clitics, as seen by (61). Note that, this is not the result of free pro-drop. Leaving out the pronoun in the second conjunct is generally disallowed, as shown by (61)c (cf. Kroeger 1993:121 fn.7).

(60) a. \* $\eta$ -anre na  $\eta$ -inu $\eta$ =a *Selayar*  
 AV-eat CONJ AV-drink=1s.A

b.  $\eta$ -anre=a na  $\eta$ -inu $\eta$ =a  
 AV-eat=1s.A CONJ AV-drink=1s.A  
 ‘I ate and drank.’

(61) a. k<um>ain at <um>inom=**ako** *Tagalog*  
 <AV.BEG>eat CONJ <AV.BEG>drink=1S.NOM  
 ‘I ate and drank.’

b. k<um>ain=**ako** at <um>inom=**ako**  
 <AV.BEG>eat=1S.NOM CONJ <AV.BEG>drink=1S.NOM  
 ‘I ate and drank.’

c. ?\*k<um>ain=**ako** at <um>inom  
 <AV.BEG>eat=1S.NOM CONJ <AV.BEG>drink

Among this small sample of languages (French, European Portuguese, South Sulawesi languages and Tagalog), all possibilities of clitic sharing and clitic type (head adjacent vs. 2P) appear to be attested. It thus remains to be shown that the proposed syntactic and prosodic correlates of clitic sharing are truly universal before we can attribute any explanatory power to the analysis of French and Standard Modern

Greek-type object markers as lexical affixes and European Portuguese and Pontic/Kozani-type object markers as  $X^0$  clitics.<sup>37</sup>

In the next section, we proceed from our discussion of clitic types and widen our scope to review various approaches to clitic placement. Here, we will focus on the division of labor between the grammatical modules in determining the placement of 2P clitics.

### 2.3 Theories of second position: 2P(rob)lems

In this section we examine the two major problems involved in 2P: that of *underlying position*,<sup>38</sup> i.e., “How and why do clitics arrive at the left edge of their domain?” and that of *non-initiality*, i.e., “Why do clitics appear to avoid the absolute edge their domain?”. These syntactic questions have taken up the lions share of discussion regarding 2P elements and it is here that we focus our attention.

The mechanisms posited for deriving left-peripheral position and non-initiality differ in almost every imaginable way. Not only has the mechanism of clitic positioning been attributed to all possible modules of the grammar, the actual source

---

<sup>37</sup> We are led to conclude that the categorial clitic properties proposed by Halpern & Fontana and C&K are epiphenomenal and not part of UG. It is possible that the apparent coincidence of positional and “ coordinational ” properties in other languages are largely an artifact of grammaticalization. The diachronic progression from 2P clitics to V-adjacent clitics to affixal agreement involves increasing integration with the verb. Nonetheless, examples of unambiguous stem coordination under affixal material is also attested, as can be seen with the Spanish adverbial suffix *-mente* in (i)-(iv) (see Artstein 2005 for coordination of word parts more generally).

- |       |   |                  |
|-------|---|------------------|
| (i)   | {decidida y atrevida}-mente ‘decisively and boldly’           | (Suñer 1975:604) |
| (ii)  | {queda y lenta}-mente ‘softly and slowly’                     | (Suñer 1975:604) |
| (iii) | {directa o indirecta}-mente ‘directly or indirectly’          | (Zagona, 1990:5) |
| (iv)  | {inteligente y profunda}-mente ‘intelligently and profoundly’ | (Zagona, 1990:5) |

It should thus not be surprising that the same could hold true for highly integrated person markers. Although grammaticalization typically operates on several properties of a lexical item simultaneously (Hopper & Traugott 1993), all aspects of integration need not proceed in lockstep and can yield some of the asymmetries discussed above.

<sup>38</sup> The term “underlying position” is unfortunately vague in referring to 2P clitics: it can refer either to the position where the full NP counterparts of clitics are generated, or to the position of 2P clitics before hypothesized phonological adjustment takes place. Here we refer to the latter meaning.

of non-initiality has also been claimed to be both purely prosodic by some (Radanović-Kocić 1996, Taylor 1996 *inter alia*) and purely syntactic by others (Rivero 1986, Jelinek 1996, Progovac 1996). In the case of more extreme analyses, 2P is seen as being entirely epiphenomenal, i.e., a result of external licensing conditions on the clitics themselves (Terzi 1999), an EPP feature attracting clitics to the left-edge (Legate 2008), or a requirement for “T-extension” in which the clitics are analyzed as T elements and [Spec,TP] can be filled by almost anything (Kucerova 2005). Bošković (2001) classifies the different approaches using the following taxonomy:

**Strong Syntax:** Clitic positioning is fully determined by the syntax

**Strong Phonology:** Phonology is fully responsible for placing clitics in second position. The syntax generates clitics in normal argument positions. They are then moved into second position as necessary in the phonology.

**Weak Syntax:** Most movement of clitics is syntactic. However, Prosodic Inversion may intervene in the phonology to satisfy the requirement of a clitic for an appropriate host.

**Weak Phonology:** Movement of clitics takes place in the syntax, and involves a considerable amount of freedom of positioning. The role of phonology is passive, filtering out certain syntactically well-formed sentences that violate phonological requirements of the clitics.

The two strong approaches are generally the most difficult to defend. Strong syntax meets with difficulties because it appears impossible to identify “second position” with a particular syntactic projection (Franks 2000:16, Progovac 2000, Bošković 1995, 2001:40). Typically, under a strong syntax analysis, clitics appear in or somewhere in the vicinity of  $C^0$  (Franks & Progovac 1994, King 1996, Progovac 1996, Roberts 1994, Schütze 1994, Tomic 1996, Wilder & Ćavar 1994) as they directly follow wh- elements in the left periphery. However, when the verb is clause initial, 2P elements will follow the verb. Strong syntax approaches are forced to argue

that in such cases V-to-C movement occurs to license the clitic (cf. Rivero 1986, Roberts 1994, Terzi 1999), but the set of elements which can license clitics in this manner is completely incoherent from a syntactic point of view and must include categories which are not typically considered mobile, e.g., adverbs (see C&K 2001 and Bošković 2001:37 for similar criticisms of Terzi 1999 and Roberts 1994, respectively). If adverbs are understood not to move freely, as is by now a common assumption, then there are unsurmountable ordering paradoxes between the position of adverbs and 2P clitic positions (see Bošković 1995, 2001:40 for details). An alternative, first suggested by Franks (1997, 1998), has 2P clitics climb to the highest functional head in their extended projection instead of always being associated with a single node. While this may do a better job of churning out the facts, it also gives the distinct impression of a non-syntactic constraint at work, as there is nothing in the theory to parallel an element which requires cyclic head-movement to an unspecified edgemost projection. More generally, under a strong syntactic approach, the ban against domain initial clitics must be phrased in completely syntactic terms. If regular syntax is able to emend phonological problems then we run into what has been termed the “look-ahead” problem in derivational frameworks: narrow syntax should not be able to forecast a prosodic violation which only occurs later in the derivation (Schutze 1994 *inter alia*).

Strong phonology appears equally unfounded as there is nothing in normal phonology which parallels the movement of words or morphemes across syntactic constituents. The most articulated strong phonology view, that of Radanović-Kocić (1996), requires that clitics are marked diacritically by a cliticization rule: “Assign the feature [+clitic] to pronouns and auxiliaries in all positions except when they carry phrasal stress or when not preceded by an unstressed element.” Another rule places these elements in their proper position: “Move all [+clitic] elements into the second

position”. These rules are of course more of a description of the facts rather than an explanation but Radanović-Kocić is credited with several important observations concerning parentheticals and appositives in relation to the role of prosody in Serbo-Croatian clitic placement.<sup>39</sup>

Thus it is the two mixed, or “weak”, approaches which have attracted the most proponents and in the interest of space, it is only these two types which we will consider in detail. In both the weak syntax and weak phonology approach, as characterized above, the phonology plays a minimal role after the derivation has positioned the clitic by regular syntax. In weak syntax, a single movement (inversion) operation is left to the phonology component which is called upon to place hostless clitics in a position where they can satisfy their prosodic requirements. In weak phonology, the phonology simply acts as a filter, without having the power to rearrange morphemes. A derivation containing a clitic without a host will simply crash as no displacement operation can take place in the phonology. The two mixed approaches sketched above still make strong predictions as to what should and should not happen with 2P clitics. In the remainder of this section, I hope to show that neither approach can handle the type of clitic patterns found most commonly with 2P clitics in Austronesian languages. In this aim, we largely follow Chung (2003) who argues convincingly against both of these approaches on the basis of the Austronesian language, Chamorro.

### 2.3.1 *Weak phonology*

The weak phonology approach, argued for by Franks (1998) and Bošković (1995, 2001) among others, crucially relies on regular syntactic operations of

---

<sup>39</sup> Notably, she argues that clitics are bound by their intonational phrase, an idea which is picked up on by Bošković (2001) in a very different type of analysis.



movement to derive clitic position. Here, the syntax gives maximum freedom to ordering clitics and phonology simply filters out derivations in which clitics have no hosts. The copy theory of movement (Chomsky 1991) treats movement as a two step process involving the copying of a “moved” element and its deletion. In this theory, multiple copies of a moved element are generated for purposes of feature checking while deletion may be triggered by orthogonal principles, such as prosodic requirements. It thus seems particularly apt for capturing 2P phenomena, as deletion can now be left to the phonological component while generation remains in syntax proper.<sup>40</sup> But the success of the copy theory of movement to handle 2P clitics is still contingent on the ability of regular syntax to produce all potential host+clitic combinations.

Progovac’s (1996) influential weak phonology analysis of Serbo-Croatian clitic placement was based on the claim that, while apparently not forming traditional syntactic constituents, all elements which can host 2P clitics correspond to elements which can be independently extracted in the syntax: “Only elements that can move to Comp or Spec of CP, or are base-generated in Comp, can support clitics” (Progovac 1996:415). She shows that certain elements which cannot be extracted syntactically but which are prosodic words also resist being broken up by 2P clitics. One example of such a constituent consists of nominal heads and following possessor phrases. The example given in (62) illustrates that clitics may not directly follow the head noun *roditelji* ‘parents’ in the complex NP *roditelji uspešnih studenata* ‘parents of

---

<sup>40</sup> Note, however, that, on this view of things, the phonology is granted a powerful mechanism for “look-back” in the attempt to preserve standard modularity and avoid look-ahead. Specifically, the stage of phonology in which the string is parsed into prosodic words must be able to identify all copies of an item erasing all but the highest copy. The null hypothesis would be that the indexes necessary for copy erasure are invisible in the phonological component, as there are no *independent phonological* processes that make reference to them. Modularity is thus only preserved at the cost of allowing more syntactic information than expected into “narrow phonology”.

successful students’. In (63), it is shown that the head is also unextractable for the purposes wh- movement.<sup>41</sup>

(62) {Roditelji[\*=su=se] uspešnih studenata} [=su=se] razišli  
 parents successful:GEN students:GEN=AUX=REFL dispersed  
 ‘The parents of the successful students dispersed.’ (Progovac 1996:418)

(63) \*Ko=su=se uspešnih studenata razišli  
 who=AUX=REFL successful:GEN students:GEN dispersed  
 (For, ‘Who of the successful students dispersed?’) (Progovac 1996:418)

Bošković (2001:15-16) further argues that the presence or absence of certain scrambling possibilities across dialects and idiolects of Serbo-Croatian correlates with the possibility of 2P clitics breaking up similar syntactic constituents. Such a picture lends considerable support to the idea that clitics and their hosts are manipulated by operations of the normal syntax.

While such an account is potentially attractive in Serbo-Croatian to the extent that the extractable constituent-clitic host correlation holds (but see fn.41), a similar account for Austronesian languages seems thoroughly hopeless as there exists a massive gap between the types of elements which can serve as clitic hosts and those which can be extracted in the normal syntax. Pre-nominal modifiers and preverbal adverbs, for instance, regularly host clitics in Tagalog but can never be scrambled outside of their immediate phrases. Chung (2003) discusses precisely this type of problem for Chamorro in great detail. To cite one of her examples, Chamorro 2P clitics, such as *hit* 1+2P.NOM, must intervene between nominal heads and following modifiers, including phrasal possessors, as shown in (64). Under a Progovacian weak

---

<sup>41</sup> These facts are highly disputed and no general agreement is found between researchers, who until recently, have relied almost solely on their intuitions. Predolac (2008) shows on the basis of corpus evidence that the clitic-host extractability generalization is unreliable. Diesing, Đurđević & Zec (2008) further show on the basis of corpus and experimental methods that the choice of first word and first phrase positioning crucially depends on pragmatics and the argument-predicate distinction, both factors having been largely ignored in previous work.

phonology approach this suggests that subextraction out of DP should be possible.

This, however, is strictly disallowed as seen by the examples in (65). Furthermore, full NP arguments can never break constituents of this type, as demonstrated by (66).

- (64) a. Famalao'an=**hit** ginin todus i islas gi Pasifika  
 women=1+2.NOM from all the islands LOC pacific  
 'We are women from all the islands of the Pacific.'
- b. Mañe'lu-ña=**hit** famalao'an si=Antonio  
 Siblings-3S.GEN=1+2.NOM women P=Antonio
- (65) a. \*Hayi [kime'=ña \_\_\_]=hit  
 who buddy=3S.GEN =1+2.NOM
- b. \*Hayi=gui' [asagua=ña \_\_\_]?  
 who=3S.NOM spouse=3S.GEN
- (66) \*Kao i ga'chóchong=ña [ädyu na taotao] si=Dolores gi bisnis?  
 QM the partner=AGR.PROG that LNK person P=D. LOC business  
 ('Is that person Dolores's partner in business?')

There is in general a very clear difference between constituents which can be intruded upon by clitics and those which can be intruded upon by non-clitic elements. Clitic positioning in such cases then appears to require a special operation which is not part of the normal syntax and thus syntactic approaches are at great odds to derive the nature of the clitic host.

### 2.3.2 *Weak syntax and Prosodic Inversion*

In the most thorough and widely cited weak syntax approach, Halpern (1995) proposed that non-constituent clitic hosts are provided by a special repair strategy, dubbed Prosodic Inversion (PI), which saves clitics from positions where they cannot satisfy their prosodic subcategorization requirement. PI takes a clitic and inverts it with an adjacent prosodic word (or hypothetically any prosodic constituent) to resolve

the clitic's prosodic dependency. The PI approach, however, still requires that clitics be positioned consistently in a particular syntactic projection. From this position, constituent hosts can be provided by topicalization or other types of movement to the left of the clitic. If nothing moves to an appropriately left-peripheral position, PI applies and offers up a prosodic word, resulting in a non-constituent host. Franks (1998), Bošković (2001), Wilder & Ćavar (1994), Ćavar (1999) and Chung (2003) among numerous others argue extensively against PI showing evidence that it both under- and overgenerates. It overgenerates in predicting that any domain initial prosodic word will be able to host clitics if clitics are still without a host after syntax. As seen above, there are several types of elements which appear to satisfy prosodic wordhood but which cannot host clitics. Among these we find the Serbo-Croatian preposition *prema*, which is claimed to be stressable but yet ineligible for clitic hosting. More troubling is the fact that the syntactic configuration which the domain initial element finds itself in very often determines whether or not it can actually host clitics. As seen above in (62), head nouns of complex NPs cannot host clitics in Serbo-Croatian. Other such non-hosts include initial conjuncts and non-case marked words contained in proper nouns (i.e. personal and place names). All such cases are an embarrassment to an approach which creates non-constituent clitic hosts by a purely phonological movement operation which should be blind to syntactic structure.

A basic problem is also seen in the disjunctive method of deriving clitic hosts either by phrasal movement above the clitic or by PI. Two examples, one from Serbo-Croatian and one from Tagalog, will serve to illustrate this point.

Anderson (1996) shows that PI appears to apply despite the topicalization of an entire phrase in Serbo-Croatian. In example (67)a, the object NP is clearly topicalized and as Halpern would predict, it hosts the clitic without inversion. But in (67)b we see that clitics may also follow the first word within a topicalized constituent and thus PI

appears to apply superfluously. Anderson also provides example (68), with a topicalized adverbial phrase, illustrating the same point.

- (67) a. Sovetske goste=je primio i predsjednik Republike Austrije Jonas  
 Soviet guests=AUX received also president republic Austria Jonas  
 ‘The president of the Republic of Austria, Mr. Jonas, also received the Soviet guests.’
- b. Sovetske=je goste primio i predsjednik Republike Austrije Jonas  
 Soviet=AUX guests received also president republic Austria Jonas  
 ‘The president of the Republic of Austria, Mr. Jonas, also received the Soviet guests.’
- (68) a. Prošle godine=su otvorili ugostiteljsku školu  
 last year=AUX opened hotel\_and\_catering school  
 ‘Last year they opened a hotel and catering school’
- b. Prošle=su godine otvorili ugostiteljsku školu  
 last=AUX year opened hotel\_and\_catering school  
 ‘Last year they opened a hotel and catering school’

The Serbo-Croatian data above shows that PI must even apply when not expected to on the basis of the syntax. Conversely, data from Tagalog shows that PI can also systematically fail to apply when expected to. Tagalog allows movement of oblique phrases and adjuncts to a clause-initial focus position. Breaking up a complex oblique phrase in this fronted position with pronominal clitics is typically judged ungrammatical (S&O 1972:187-193, Kroeger 1993). Rather, clitics must follow the entire focus phrase as shown in (69). This stands in contrast to the situation obtaining with regular prepositional phrase predicates which are commonly broken up by pronominal clitics, as shown in (70).

- (69) Sa=ma-laki[\*=**ako**]=η syudad[=**ako**] naka-tira[=\***ako**]  
 OBL=ADJ-big=1S.NOM =LNK city STA-stay  
 ‘I live in a big bity’
- (70) Sa=ma-laki[=**ako**]=η syudad[<sup>?</sup>=**ako**]  
 OBL=ADJ-big=1S.NOM =LNK city  
 ‘I’m in a big city’



with a following modifier in (73).

- (72) Moja=**je** sestra stigala  
my=**AUX** sister arrived  
'My sister arrived.'
- (73) Studenti[**\*=su**] iz Beograda[=**su**] upravo stigli  
students=**AUX** from Beograd just arrived  
'Students from Beograd have just arrived'

Halpern tentatively proposes the two principles in (74) and (75) to handle such cases.

- (74) A clitic must be contained in the same phonological phrase as its host
- (75) The left edge of the head of a branching constituent corresponds to the left edge of a prosodic phrase

While (74) is a standard assumption and falls out independently from Strict Layer Hypothesis (Selkirk 1984), employing such a principle for the placement of 2P clitics as envisioned above by Halpern is not entirely straightforward. In a derivational framework, this would require an additional cycle specifically for the placement and parsing of 2P clitics. Specifically, prosodic phrases must be parsed before the phonology countercyclically resolves a dependency on the prosodic word level. More problematically, it is not clear how the above principles derive the correct post-phrasal ordering, as in the case of (73). The parsing algorithm in (75) yields the following parses for (72) and (73) at the end of the syntactic derivation (with the clitics in their base generated position in the left periphery prior to PI).

- (76) =**su** PPh[**prijatelji** moje PPh[**sestre** stigli  
=**AUX** friends my sister arrived
- (77) =**je** moja PPh[**sestra** stigala  
=**AUX** my sister arrived

In accordance with (75), prosodic phrase edges are aligned to left edges of syntactic heads of branching constituents. This has the effect of breaking up branching syntactic phrases into multiple prosodic phrases when a modifier precedes the head. Because the clause in (76) begins with a head of a branching phrase, a prosodic phrase boundary separates the base position of the clitic from the following prosodic word. In the case of (77), on the other hand, no prosodic phrase boundary separates the clitic from the following prosodic word and thus enclisis to that word is permissible.

One serious problem which is left undiscussed by Halpern is how the clitic in (76) finds the correct host given the above principles. Presumably, (74) simply disallows PI to take place as the attested placement after *sestre* would also violate the same principle. This could lead the derivation to crash and only allow variants which moved the entire host phrase to the left periphery above the base position of the clitic. However, this would then not differ significantly from approaches which allow true syntactic movement to be triggered by a phonological dependency (i.e. the problem of look-ahead discussed above).

Empirically, PI has been seen from the above to both under- and overgenerate. On the theoretical side, there is neither a principle nor precedent for base generating all the clitic elements in the leftmost position based on their function. Aspectual markers, pronominals, and mood and focus clitics must all be generated in the same leftmost position and this position must be outside left peripheral focus projections, as in (71). While such a position may look appropriate for an element such as a question marker, it would be most unusual for a phrase hosting pronominal arguments. Uglier yet is the fact that oblique pronouns in Tagalog, for which it would be highly infelicitous to posit an agreement projection, can also optionally appear in clitic position. Finally, as originally noted by Perlmutter (1971), a syntactic account must still invoke additional mechanisms for ordering clitics in relation to each other as



ordering relations within the clitic cluster can be highly irregular from a syntactic perspective (i.e. based on syllable count as in Tagalog). Regular syntax as the primary means of clitic placement can thus be shown to fail for Tagalog and other Philippine languages. The arguments above apply equally to the weak and strong syntactic approaches.

### 2.3.3 *Distributed Morphology and Morphological Merger*

The theory of Distributed Morphology (DM) (Halle & Marantz 1993, 1994, Embick & Noyer 1999/2001) envisions a rather different division of labor between the grammatical modules. One of the pillars of DM is the creation of all morphological structure through syntax, known as “syntax all the way down” (Harley & Noyer 1999). But “syntax” in the sense of DM is not precisely syntax as envisioned by previous theories. Rather, the emphasis on syntactic derivation is meant to stand in contrast to lexicalist approaches where word formation is undertaken prior to syntax in a dedicated lexical component. Within DM, word formation (in addition to many cases of clisis) is handled post-syntactically by many of the same mechanisms operative in the regular syntax. Syntax manipulates features which are later filled in by “Vocabulary Items” (roots, affixes, clitics, etc.) in the post-syntactic morphological component (making DM a “Late Insertion” model of morphology). The distinction between affixation and clisis, which often represented a difference between lexical derivation and syntactic derivation in lexicalist theories, is treated as epiphenomenal in DM. This also does away with many of the claims of correlations between the syntactic and phonological behavior of clitics, which, as we saw earlier, were difficult to uphold cross-linguistically. However, the substantive claim of “syntax all the way down” is softened by the reliance on several heterogeneous post-syntactic operations

which operate solely on affixes and clitics in the guise of functional categories.<sup>43</sup> For this reason, locating a DM approach to 2P phenomena in Bošković's taxonomy is not entirely straightforward as the conception of syntax is far wider than it is in other theories. Nonetheless, we examine here what DM has to say about 2P with the negative conclusion that it offers no aid to the problems outlined in the previous subsection.

Marantz (1988) and Embick & Noyer (1999/2001) (herein E&N) propose a theory of clisis and general dislocation which makes crucial reference to both structural (syntactic) relations and linear relations. Regarding whether or not all clitic positioning should be handled by syntax proper, E&N take the position that the constraints of narrow syntax can, in principle, be loosened to handle any ordering relations but that it would be a mistake to do so. Rather, clitic positioning which is not derived by syntax proper can be handled by any of three post-syntactic operations: Lowering, Local Dislocation, and Prosodic Inversion (the last of which the authors are less than fully committed to). These operations are meant to manifest movement at different stages of the derivation and therefore make reference to very different types

---

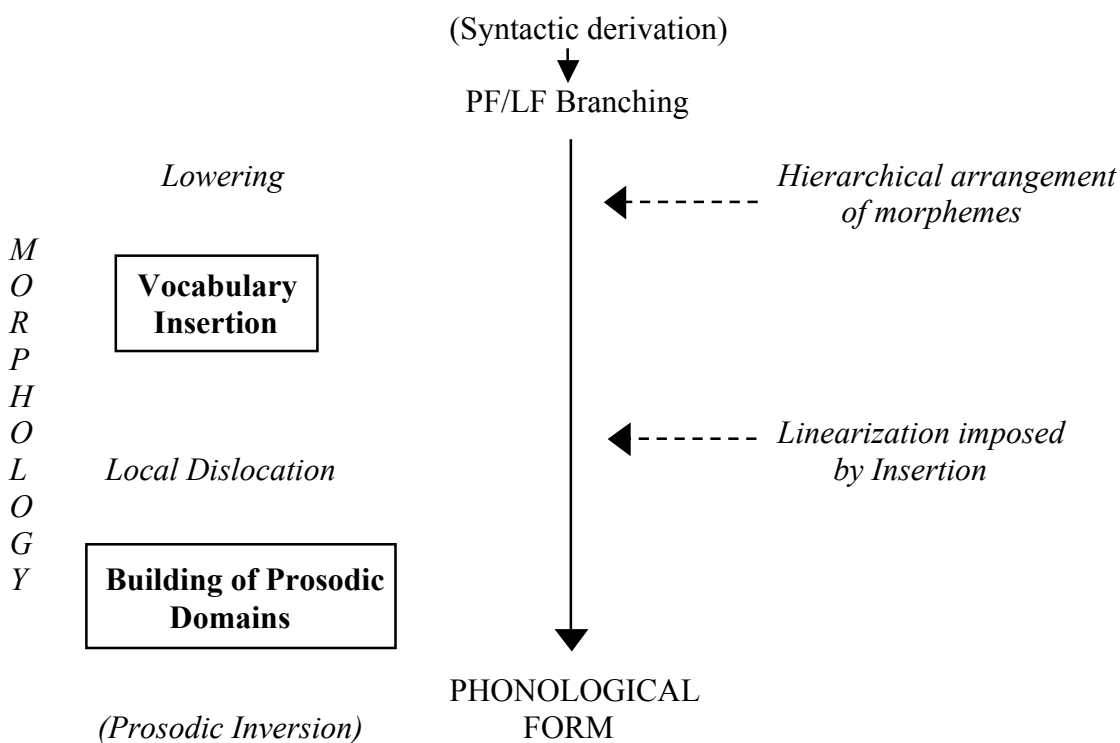
<sup>43</sup> This point has also been noted by several critics of the approach, e.g., Baker (2003) who sticks closer to the view of Chomsky (1970):

“This suggests that Halle and Marantz (1994) go a bit too far in saying that the internal morphological structure of words is ‘syntax all the way down.’ ...[O]nce the syntactically predictable morphology has been stripped away, there remains a residue of morphology that seems to have nothing to do with syntax. This residue includes a rather wide range of not-very-productive and semantically idiosyncratic derivational morphology, as well as root compounding and those language-particular aspects of inflection that revolve around grammatical gender, concord, and purely formal matters of inflection such as the Indo-European theme vowels and the Mohawk noun suffixes. There is perhaps a generative morphology of quite modest power after all, distinct from syntax, that deals with the internal structure of these linguistic objects. I have no reason to be dogmatic on this point; if good reasons come to light for saying that the adjective *foggy* is formed in the syntax, so much the better. For the time being, however, complicating the syntax with derivations of this kind seems likely to do more harm than good.” (Baker 2003:280)

of representations. All three reordering operations make use of the notion of Morphological Merger, as defined by Marantz (1988):

- (78) Morphological Merger  
 At any level of syntactic analysis (D-Structure, S-Structure, phonological structure), a relation between X and Y may be replaced by (expressed by) the affixation of the lexical head of X to the lexical head of Y. (Marantz 1988:261)

The set of operations taking place on the PF side of the derivation after syntax branches off into PF and LF (E&N:558) can be seen explicitly in the left hand of Figure 2.1 below, which illustrates the architecture of DM according to E&N (1999/2001)



**Figure 2.1.** DM architecture (Embick & Noyer 1999:273)

Crucially, certain locality constraints must be respected at all stages of the derivation but the nature of the locality involved changes with the differing types of

representations as the derivation proceeds. Lowering takes place at the earliest point in PF and thus makes reference to syntactic structure before the point of Vocabulary Insertion. At this point in the derivation, X and Y refer to syntactic phrases in hierarchical relation to each other. By Lowering, the head of a dominating phrase X can attach to the head of its complement phrase Y regardless of intervening material (e.g., adverbs and adjuncts).

$$(79) \text{ Lowering of } X_0 \text{ to } Y_0 \\ \text{XP}[X_0 \dots \text{YP}[\dots Y_0 \dots]] \rightarrow \text{XP}[\dots \text{YP}[\dots Y_0[Y_0 + X_0] \dots]]$$

The ability to skip intervening syntactic material is meant to be responsible for the positioning pattern displayed by clitics such as the Bulgarian definite article. This clitic attaches to the head noun in a simple NP such as (80)a, but attaches to a preceding adjective when present, as in (80)b. However, when the adjective itself is modified, as in (81), the clitic can only attach to the adjective and not the adverb.

- |      |   |  |
|------|---|--|
| (80) | a. kniga= <b>ta</b><br>book=DEF                         | b. xubava= <b>ta</b> kniga<br>nice=DEF book  |
| (81) | a. *mnog= <b>at</b> star teatər<br>very=DEF old theater | b. mnogo starij= <b>ə</b> teatər<br>very old=DEF theater<br>'the very old theater' |

Syntactic constituency must be relevant to the positioning of the clitic in (80) and (81), otherwise adjectives and adverbs would be treated on par with each other. This is captured by Lowering as hierarchical relations are still visible, thereby allowing reference to notions such as “head of the initial phrase”.<sup>44</sup>

---

<sup>44</sup> Note that this analysis is highly dependent on how adjectives and adverbs are positioned in DP structure (with adjectives heading their own phrases and adverbs being adjuncts). Other complications for the Bulgarian data are discussed in chap. 5.

After Vocabulary Insertion, hierarchical relations are traded in for linear adjacency relations. At this point, reordering by Morphological Merger is sensitive to all intervening syntactic material because linear adjacency must be preserved from the input to the output. However, adjacency relations are still not defined over a string of atomic elements but over constituents. Although E&N treat adjacency and precedence as relations holding on the same level of representation, we can also think of adjacency as the relation relevant to the structure before the building of prosodic domains and precedence as the relevant relation after this process has taken place. Modifying E&N's notation by taking " $\Leftarrow$ " to represent the adjacency relation after Vocabulary Insertion, we can see the potential outputs of (82)a.<sup>45</sup>

- |      |                                      |               |                      |               |
|------|--------------------------------------|---------------|----------------------|---------------|
| (82) | Adjacency relations                  | $\rightarrow$ | Precedence relations |               |
|      | a. $[X \Leftarrow [W \Leftarrow Y]]$ |               | b. $[X W Y]$         | c. $[X Y W]$  |
|      |                                      |               | d. $[W Y X]$         | e. $[Y W X]$  |
|      |                                      |               | f. $*[W X Y]$        | g. $*[Y X W]$ |

If Morphological Merger does not apply (82)a will be spelled out as (82)b. However, Morphological Merger can, in theory, apply to any of the elements within the string to also yield (82)c-e which all preserve the relations in (82)a. Those outputs which do not respect the adjacency relation within the constituent  $[W \Leftarrow Y]$  can be immediately ruled out. In this sense, dislocation is considered to be local.

Note, however, that the possibility of creating segments via adjunction could radically weaken the predictions of the adjacency requirement, in particular, allowing orders like (82)f-g. Concretely, if X were allowed to adjoin as a segment of W as in (83)b, and X could still be considered adjacent to the constituent  $[W \Leftarrow Y]$  from its embedded position, then locality would be far less constrained. Confusingly, although

---

<sup>45</sup> Marantz (1988, 1989) and E&N employ "\*" to represent adjacency and "+" to represent the precedence which follows from adjunction.

hierarchical relations are erased at this stage, the transformation in (83)b is deemed by E&N as permissible because “both [ $\Leftrightarrow$ ] relations in [(83)a] have been either respected or properly converted by Local Dislocation” (E&N 2001:563).

- (83) Adjacency relations  $\rightarrow$  Precedence relations  
 a.  $[X \Leftrightarrow [Z \Leftrightarrow Y]]$                       b.  $[_z[Z X] Y]$

The answer to how the adjacency relations in (83)a are actually maintained in the precedence relations in (83)b is found in a footnote: “The affixation of X to Y might first involve rebracketing under adjacency, such that  $[X \Leftrightarrow [Z \Leftrightarrow Y]]$  becomes  $[[X \Leftrightarrow Z] \Leftrightarrow Y]$  prior to inversion of X” (E&N 2001:563 fn.9). But the possibility of free rebracketing feeding inversion clearly expands the power of dislocation operations in DM and without more explicit constraints, it is no longer clear what the predictions are. The idea, suggested in the above citation that rebracketing may feed inversion also seems to contradict the stated predictions of the theory. E&N claim that, “if X is an element peripheral in some constituent C, X will not be able to invert with an element Y that is outside of the constituent C [(84)b], although leaning is possible [(84)c].” In (84)c, leaning is claimed to be possible through rebracketing. However, if rebracketing can feed inversion then clearly (84)b should also be permissible.

- (84) a.  $[ \dots Y ] \Leftrightarrow_c [X \Leftrightarrow Z]$   
 b.  $[ \dots X Y ] \Leftrightarrow_c [Z]$  impossible inversion  
 c.  $[ \dots Y X ] \Leftrightarrow_c [Z]$  possible leaning

We can now ask more concretely at this juncture what the predictions of Morphological Merger are regarding 2P clisis. Fortunately, E&N spell this out explicitly:

“The theory predicts certain interactions concerning domains that are accessible for Merger operations and concerning the relative ordering

of operations. First, a complex  $X_0$  created in syntax (or by Lowering) cannot be infixes within another  $X_0$  during Morphology. In other words, ‘second’ position for an MWd is after (or before) the first (last) MWd in a phrase and nowhere else.” (E&N 2001:577)

As discussed above for Halpern’s theory of PI, one of the primary challenge of capturing clitic placement in Tagalog is the proper derivation of impenetrable constituents. The existence of impenetrable constituents larger than the word is a fact not only of Tagalog and Philippine languages but of all languages which display 2P phenomena. Regarding impenetrability and post-phrasal positioning, E&N go on to comment:

“Klavans (1995) and Halpern (1992b) (among others) discuss cases in which second position appears to be either after the first word or after the first phrase. On the present proposal, positioning after the first phrase cannot arise from Local Dislocation. It must be the case that the initial XP in question has raised (e.g., by a topicalization fronting) to sentence-initial position. Whether or not a prosodic operation is needed in addition to movement for cases in which an XP precedes a clitic is an open question.” (E&N 2001:577 fn.29)

As seen from the above, E&N do not depart from the Halpernian approach of treating post-phrasal positioning as resulting from movement of an XP higher than the base position of the clitic. Thus, in addition to making unclear prediction regarding “invertable” elements, the same criticisms which apply to Halpern’s XP raising analysis in regard to impenetrability also apply to E&N’s approach.

#### 2.3.4 *Attachment to phonological hosts and Prosodic Subcategorization*

Having now presented arguments against every type of approach listed above, what is left? Chung (2003) argues that only a prosodic approach can handle the positioning of weak pronominal clitics in Chamorro. The close correspondence between Chamorro facts and their Tagalog analogues may suggest a similar approach

is appropriate for Philippine clitic systems. A careful review of Chung’s analysis is thus in order.

Chamorro has a set of independent and weak pronouns shown in Table 2.4. The weak pronouns are in many cases reduced forms of the independent forms but, more importantly, they differ from their independent counterparts in their positioning. Independent pronouns have the same basic distribution as full DPs while weak pronouns are positioned as 2P elements.

**Table 2.4.** Chamorro Pronouns

	<i>Weak</i>	<i>Independent</i>
1S	yu’	guahu
2S	hao	hagu
3S anim.	gui’	guiya
1P	hit	hita
1+2P	häm	hämi
2P	hämyu	hämyu
3P anim.	siha	siha

As previewed earlier, Chung convincingly dispatches several varieties of analyses which rely on syntactic movement for clitic positioning. Firstly, full DPs are prohibited from occupying common 2P positions, as seen from (85), where the weak pronominal interrupts a complex DP containing a possessor in (85)a.

- (85) a. Kao patgon-ña=**hao** ädyu na ma’estra? *Chamorro*  
 QM child-3S.GEN=2S.NOM that LNK teacher  
 ‘Are you the child of that teacher?’
- b. \*Kao patgon-ña **si=Dolores** ädyu na ma’estra?  
 QM child-3S.GEN P=Dolores that LNK teacher  
 (For, ‘Is Dolores the child of that teacher?’)

Secondly, no movement operation in Chamorro could possibly extract non-constituent clitic hosts, as shown in (86), where an NP is unsuccessfully subextracted from a DP containing a possessor and (87), where the possessor itself is



unsuccessfully extracted. Subconstituents of DP must, however serve as clitic hosts, as seen in examples such as (88).

- (86) a. ?\*[Mismu kareta-nñiha]=ha' ma-sugun i famalao'an *Chamorro*  
 own car-3P.GEN=EMPH PASS-drive the women  
 (For, 'The women's own car was driven.')
- b. ??[Patgon-ña] māmaigu' i palao'an.  
 child-3S.GEN sleep:PROG the woman  
 (For, 'The woman's child was sleeping.')
- c. \*Kao i ga'chóchong=ña [ādyu na taotao] si=Dolores gi bisnis?  
 QM the partner=AGR.PROG that LNK person P=D. LOC business  
 ('Is that person Dolores's partner in business?')
- (87) a. \*Hayi [kime'=ña \_\_\_]=hit  
 who buddy=3S.GEN =1+2.NOM
- b. \*Hayi=gui' [asagua=ña \_\_\_]?  
 who=3S.NOM spouse=3S.GEN
- (88) a. Famalao'an=**hit** ginin todus i islas gi Pasifika  
 women=1+2.NOM from all the islands LOC pacific  
 'We are women from all the islands of the Pacific.'
- b. Mañe'lu-ña=**hit** famalao'an si=Antonio  
 siblings-3S.GEN=1+2.NOM women P=Antonio  
 'We are Antonio's sisters.'

The ability to host clitics without being able to undergo syntactic extraction holds not only for complex DPs of the type above but also for nouns preceded by most types of modifiers. Just as in Philippine languages, modifiers are joined to the elements they modify by a linker (*Chamorro na*). As seen in (89), modifiers of NP cannot be extracted by the regular syntax, yet they regularly host clitics, as in (90).

- (89) a. I más amku' na chi'lu-hu palao'an k<um>ékuentus  
the most old LNK sibling-1S.GEN female <AV>speak:PROG  
  
yan i principal.  
with the principal  
'My oldest sister was talking to the principal.'
- b. \*I más amku' k<um>ékuentus na chi'lu-hu palao'an  
the most old <AV>speak:PROG LNK sibling-1S.GEN female  
  
yan i principal.  
with the principal  
(For, 'My oldest sister was talking to the principal.')
- (90) a. Bunitu=**gui'** na lahi pa'gu.  
handsome=3S.NOM LNK man now  
'He was a good-looking man now.'
- b. Más yā-hu=**hao** na taotao.  
most like-1S.GEN=2S.NOM LNK person  
'You're the person I like most.'
- c. I más amku'=**gui'** na chi'lu-hu palao'an.  
the most old=3S.NOM LNK sibling-1S.GEN female  
'She's my oldest sister.'

Finally, it cannot be the case that Prosodic Inversion is responsible for non-constituent hosts because the host is often larger than a prosodic word. This can be seen by comparing adverbial clitics, which do appear roughly in the position following the first prosodic word in their domain, with weak pronominal clitics, which may take a larger host, as seen in (91). Thus, if Prosodic Inversion were responsible for placing clitics in Chamorro it would have to be parametrized for different prosodic constituents depending on the type of clitic in question, a situation which was explicitly predicted to be impossible by Halpern (1995).

- (91) Fiu=**ha'** man-malagu=**häm** *Chamorro*  
often=EMPH PL-run=1P.NOM  
'We very often run.' (Chung 2003:567)

Another challenge of the Chamorro pattern for all accounts is a split in positioning possibilities within predicates according to their lexical category. Weak pronominals cannot appear after complements of verbal and adjectival predicates but can follow entire nominal and prepositional predicates.<sup>46</sup>

- (92) a. Águaguat na patgun=gui' *Chamorro*  
naughty LNK child=3S.NOM  
'He's a naughty child.'
- b. Más yä-hu na taotao=hao  
most like-1S.GEN LNK person=2S.NOM  
'You're the person I like most.'
- c. Man-sueti=n taotao=hit  
PL-lucky=LNK people=1+2P.NOM  
'We are such lucky people.'
- (93) a. Pära manu=yu' pa'gu? *Chamorro*  
to where=1S.NOM now  
'Where do I [go] now?' (Cooreman 1983:82)
- b. Ginin San Roque na songsung=yu'  
from San Roque LNK village=1S.NOM  
'I'm from San Roque village.'
- c. Hu-faisin kao ginin i guälu'=ha'=gui' magi.  
1S-ask QM from the farm=EMPH=3S.NOM to.here  
'I asked [him] if he had just [come] from the farm.'

Chung's analysis understands Chamorro clitics as being positioned according to a Prosodic Subcategorization frame (Inkelas 1990) as given in (94).

(94) i[p[p[ ]\_ ]]

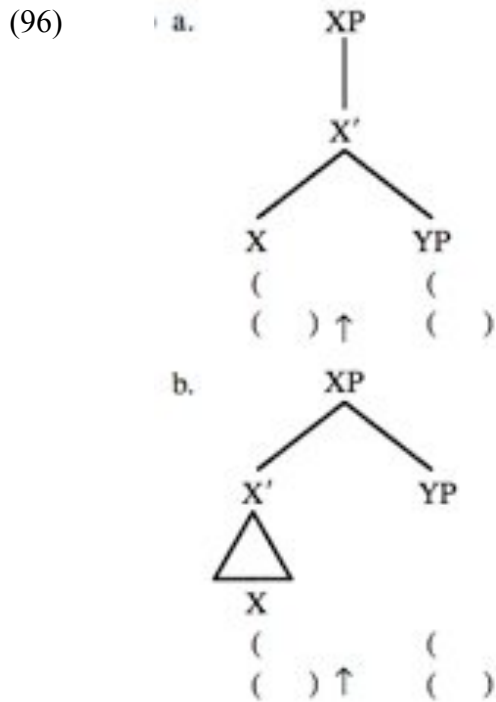
<sup>46</sup> Diesing, Đurđević & Zec (2008) show that this pattern holds as a strong tendency in Serbo-Croatian. A similar pattern has also been claimed to exist in Tagalog by Sityar (1989) (see also Kroeger 1992) although it will be shown that the facts have been overstated there.

What (94) states is that weak pronominals attach to the leftmost prosodic phrase within the intonational phrase to form a recursive prosodic phrase including both the initial phrase and the clitic. In order for such an approach to go be successful, it is of course necessary to elaborate a theory of how syntactic structure gets mapped to prosodic structure. The prosodic mapping algorithm in (95) refers to phrase structural constituents across across categories.

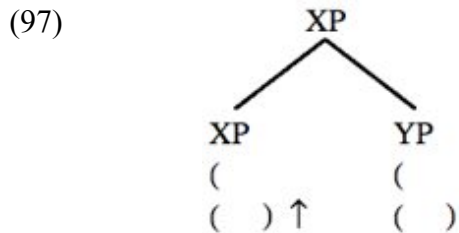
(95)      Align the left edge of XP with the left edge of a p-phrase.

As in other work, this mapping algorithm is claimed to skip functional projections and optionally skip adjectival and adverbial modifiers, i.e., it does not treat them as their own maximal projections. Chung (2003:573) goes on to argue that languages differ parametrically in regarding to how adjuncts are treated by mapping algorithms such as (95), in particular whether or not they are included within the XPs which they adjoin to. Based on the syntactic evidence discussed in Chung (1998:325-332), Chamorro appears to treat adjuncts as optionally included within their containing XPs. Chung argues that this optionality may help explain the variation found with weak pronominal placement in modificational structures, among others.

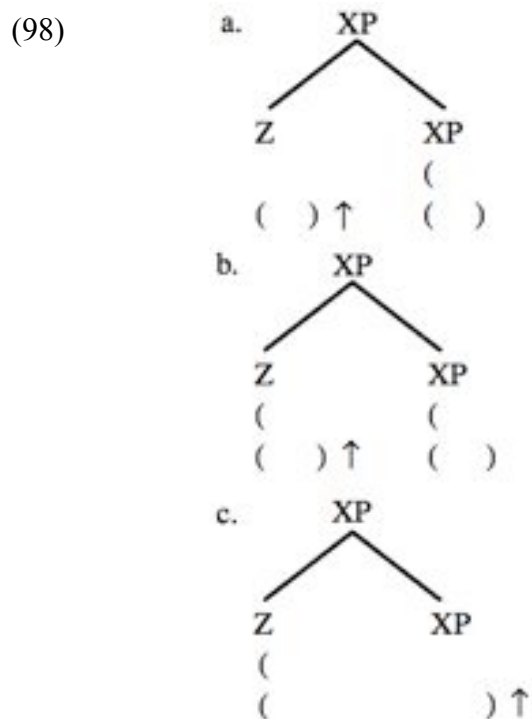
To make matters more concrete, observe the structure in (96)a, which shows the parsing of an XP into prosodic phrases and the placement of weak pronominals in a typical head-complement structure. The structure in (96)b illustrates the same information with a rightwards specifier. In both of these cases, the parsing algorithm in (95) looks at the left edges of XPs and aligns it to a prosodic phrase boundary. After the right edges are filled in by default, weak pronominals are positioned after the initial prosodic phrase following (94), as indicated by the arrow.



Following classic assumptions concerning X-bar theory, complements and specifiers differ from adjuncts in that they do not create multiple segments of the immediately dominating node. The creation of multiple segments by adjunction potentially effects the parsing algorithm in an interesting way. In the case of right-adjunction of YP to XP, shown in (97), we find a structure which does not differ significantly from the examples in (96) for the purposes of the mapping to prosodic structure. In neither case does the algorithm encounter a left edge of an XP which is not also the left edge of the maximal XP ( $XP^{\max}$ ).



With left-adjunction, however, the algorithm does encounter a left edge of an intermediate, non-maximal XP. Recall that adjuncts are optionally included within the phrases which they adjoin to. Assuming that adjuncts are parsed separately, if the left adjunct is not phrasal, then the algorithm will only see the left edge of the following phrase, i.e., the lower segment, as illustrated in (98)a. However, the excluded material will be parsed into a prosodic phrase to satisfy exhaustivity, and clitics will be placed after this initial phrase as shown in the bottom line of (98)a. If the left adjunct is phrasal then both the adjunct and the lower segment of the category adjoined to will be parsed into prosodic phrases by the algorithm in (95), as shown in (98)b. Optionality arises from the situation illustrated in (98)c, where the left adjunct is phrasal but taken to be included within the containing XP. In this case, the parsing algorithm will not “see” the left edge of the lower segment and the adjunct will be consequently phrased together with following material. Weak pronominals will then follow the larger structure, as indicated by the arrows.



This analysis correctly predicts that the optionality found with clitic placement in ADJ + NOUN constituents should disappear when the left adjunct contains an embedded phrase. This is because the left edge of the embedded phrase will be aligned with the left edge of a new prosodic phrase and, assuming no recursivity, will close off the preceding prosodic phrase. The preceding material once closed off into a prosodic phrase will obligatorily host weak pronominals because it will now satisfy their prosodic subcategorization requirement. This then predicts the optionality in (99) with the lack of optionality in (100) by the phrasal status of the embedded *ki guahu* ‘than me’.

(99) Bunitu[=gui’]        na lahi[=gui’] pa’gu  
 handsome=3S.NOM LNK man        now  
 ‘He was a good-looking man now.’

(100) Metgot-ña[=hao]        ki guahu[\*=hao]  
 strong-COMPAR=2S.NOM than 1S.INDEP  
 ‘You are stronger than me.’

Chung’s approach to clitic placement in Chamorro is far more applicable to Tagalog than any of the previous theories reviewed. It is additionally both elegant and able to account for the observed variation in a principled manner. However, there are still certain fundamental difficulties with applying it to Tagalog which will ultimately force us to seek another solution.

The main problem encountered in Tagalog but not, apparently, in Chamorro is the existence of a class of impenetrable constituents which can only be described in syntactic terms. Tagalog differs from Chamorro in allowing movement of oblique phrases and certain adjuncts to a focus position at the left edge of the clause. Phrases which are moved to this position, regardless of their size, cannot be intruded upon by clitics from the main clause. Compare (101), with a focus fronted oblique phrase, and (102) with an oblique phrase predicate (repeated from (69) and (70)). In the former, a

pronominal clitic associated with the predicate cannot intrude into the fronted phrase, but in the latter, the predicate oblique is preferably split by the subject clitic.

(101) Sa=ma-laki[\*=**ako**]=ŋ      syudad[=**ako**] naka-tira[=\***ako**]  
 OBL=ADJ-big=1S.NOM =LNK city                      STA-stay  
 ‘I live in a big bity’

(102) (Ná)sa=ma-laki[=**ako**]=ŋ      syudad[<sup>?</sup>=**ako**]  
 OBL=ADJ-big=1S.NOM =LNK city  
 ‘I’m in a big city’

Treating the focus phrase as an adjunct does not help us here, as this could only render the left edge of the VP invisible for the parsing algorithm and result in placing the clitics in the final position in (101), an undesired outcome. What is necessary is a rule which “flattens” the focused oblique into a single prosodic phrase without altering the following material. The focus phrase will then constitute the first prosodic phrase in the intonation phrase in a sentence such as (101), thus deriving the correct position. But such an analysis for Tagalog is forced to rely on the same prosodic subcategorization frame found in Chamorro, where weak pronominals follow the first prosodic phrase. But unlike Chamorro, Tagalog clitics generally appear after the first prosodic word in their domain. Recall from (71) above, repeated below as (103), that, when negation precedes the focus phrase, clitics generally follow negation directly, despite negation being unable to constitute a prosodic phrase on its own.

(103) Hindí[=**ako**] sa=ma-laki=ŋ      syudad[<sup>?</sup>=**ako**] naka-tira[=\***ako**]  
 NEG=1S.GEN OBL=ADJ-big=LNK city                      AV.BEG-study  
 ‘I don’t live in a big bity’

While both Chamorro and Tagalog show a certain disregard for syntax when determining clitic hosts in the general case, only Chamorro appears to allow interruption of all types of syntactic constituents. For instance, in (104), a weak



pronominal interrupts a complex interrogative in [Spec,CP]. This is often judged ungrammatical in Tagalog, as can be seen from comparing the attested (105)a with the variant in (105)b.

- (104) Hayi=**hao** na famalao'an gāi-che'lu?  
 who=2S.NOM LNK women have-sibling  
 'Which girls have you as a sibling?'
- (105) a. Saan panadéro=**kayo** b<um>í~bili naŋ=tinápáy?  
 where:LNK baker=2P.NOM <AV.BEG>INCM~buy GEN=bread  
 'From which baker do you buy bread?'<sup>47</sup>
- b. \*Saan=**kayo**=ŋ panadéro b<um>í~bili naŋ=tinápáy?  
 where=2P.NOM=LNK baker <AV>INCM~buy GEN=bread

Interestingly, it is not the case that focus fronted constituents cannot contain any clitics at all. Rather, they are only prevented from containing clitics associated with an element external to the focus phrase itself. Adverbial and genitive clitics which are associated with elements in the focus phrase must appear within it, as seen in (106), where the focus phrase contains two such adverbials.

- (106) Saan=**pa=ba**=ŋ panadéro=kayo b<um>í~bili naŋ=tinápáy?  
 where=still=QM=LNK baker=2P.NOM <AV>INCM~buy GEN=bread  
 'From which other baker do you buy bread?'

A prosody based account would be at pains to describe these facts. First of all, there is a degree of syntactic sensitivity which cannot be easily be captured using standard mapping algorithms. Second of all it would require a rather complex derivational process whereby adverbial clitics are placed following the first prosodic word in the focus phrase after which the prosodic structure of this phrase is flattened (so that internal prosodic boundaries are invisible) and external clitics are added. This

<sup>47</sup> From: [www.yehey.com/boards/default.aspx?g=posts&m=250522](http://www.yehey.com/boards/default.aspx?g=posts&m=250522)

approach seems to mask the fact that Tagalog pronominal clitics have to be in a particular surface syntax relation with the predicates they associate with. This relationship is not entirely unlike the c-command relation; c-command of the predicate from within the focus phrase would be impossible and it is this which renders the pronominal positioning in (101) and (105)b ungrammatical.

In the next section, we introduce Optimality Theory approaches and develop a slightly different type of prosodic analysis for 2P phenomena in Tagalog.

### 2.3.5 *Optimality Theory approaches to 2P*

Descriptively speaking, the positioning of clitics after the first available host can be seen as the result of two conflicting constraints: 2P clitics want (i) to be spelled out as early as possible but (ii) cannot appear initially within a certain domain. Following the intuition that 2P clitics are borne of this simple conflict rather than regular syntactic mechanisms, Optimality Theoretic (henceforth OT) approaches (Prince & Smolensky 1993) seek to account for this by framing 2P as an optimal satisfaction of two violable constraints, one which aligns clitics to the edge of a syntactic constituent, e.g. LEFTMOST/ALIGN-L (cl,TP), and one which prevents clitics from surfacing initially in a syntactic or prosodic domain, e.g. NON-INITIAL (cl, CP/PPh). The relative ranking of these constraints will determine whether a given clitic will surface on the edge of the constituent which it selects or in 2P within that constituent. The former will be derived if NON-INITIAL dominates LEFTMOST and the latter, if LEFTMOST dominates NON-INITIAL. The mechanics of this analysis will be demonstrated by way of a comparison with the well known OT analysis of Tagalog clitics.

The pioneering work on clitics in the OT framework carried out by Anderson (1996 *et seq*) and Legendre (1998 *et seq*) assumed Anderson's (1992) theory of

morphology in its approach to what kinds of elements clitics are underlyingly. For Anderson, clitics form a natural class with affixes as opposed to word-level syntactic atoms and are accordingly referred to as “phrasal affixes”. One piece of evidence for the clitic-affix connection discussed by Anderson (1992) is the analogy between 2P clisis on the phrasal level and infixation on the word level. In both cases, the element in question appears to avoid initial position within its domain. According to the Prince & Smolensky’s (1993) now standard account, infixes are required by principles of morphological concatenation to appear aligned to one of the edges of their stem but are displaced to a stem internal position because of phonological difficulties. Under this view, in both infixation and 2P clisis, a situation is avoided in which a prosodic domain begins with an infelicitous initial element by sacrificing edge alignment. Prince & Smolensky’s (1993) analysis of Tagalog infixation can be profitably compared with that of 2P clisis in the same language. Tagalog affixes can be divided into two categories on the basis of their phonological shape. The first category, including affixes such as /pag/, /mag/, /ka/, /pa/ are C(onsonant)-initial while the second category, including /in/ and /um/ are V(owel)-initial. The “place-shape generalization”, first noted by Anderson (1972) (see also Cohn 1992, Kaufman 2003), states that the locus of affixation is predictable by the basic shape of an affix. The V-initial affixes infix while the C-initial affixes prefix to the edge of their host, not only in Tagalog but across Austronesian languages. The infixation of V-initial affixes makes good sense in light of what is argued to be the universally optimal/unmarked nature of CV syllables (Venneman 1988). The vowel of the affix is provided with an onset by the initial consonant of the stem, and the final consonant of the affix syllabifies in turn as an onset for the following root vowel, as shown schematically in (107)a. When the stem is vowel initial, the affix is observed to simply prefix, as infixation in this case would not confer any improvements. This is shown in (107)b.

- (107) a. [C<VC>V...]<sub>STEM</sub> b. VC-[V...]<sub>STEM</sub>

The observation that non-canonical morpheme concatenation appears to satisfy universal prosodic preferences is captured directly by positing two conflicting constraints, one inherently morphological ALIGN (AFF, STEM), and the other inherently phonological ONSET, or alternatively, NO CODA. If we assume, as we must in parallel OT, that phonological constraints can dominate morphological constraints, then the ranking of ONSET above ALIGN-L, predicts that the affix will be displaced from the edge of the stem when it can gain an onset by infixation. The difference in the affix-stem interactions with C-initial and V-initial roots is shown in Tableau 2.1 and Tableau 2.2, respectively.<sup>48</sup>

root: <i>kagat</i> ‘to bite’ affix: <i>um</i>	ONSET	ALIGN-L (AFF, ROOT)
a. <u>um</u> kagat	*!	
b. <sup>☞</sup> k <u>um</u> agat		*
c. ka <u>um</u> gat	*!	**
d. kag <u>um</u> at		***!

**Tableau 2.1.** Infix with C-initial stem (Prince & Smolensky 1993)

root: <i>ábot</i> ‘to reach’ affix: <i>um</i>	ONSET	ALIGN-L (AFF, ROOT)
a. <sup>☞</sup> <u>um</u> abot	*	
b. a <u>um</u> bot	**!	*
c. ab <u>um</u> ot	*	***!

**Tableau 2.2.** Infix with V-initial stem (Prince & Smolensky 1993)

The idea which seems readily extendible to 2P clisis is the insight that prosody can outrank morphosyntax in order to displace an edge-aligned morpheme. In Tableau 2.3, we see an Anderson’s analysis of how the Tagalog pronominal clitic *sila* is positioned in 2P.

<sup>48</sup> This represents somewhat of a simplification of the facts, see Yu (2007), McCarthy (2003) and Kaufman (2003) for discussion and revisions to Prince & Smolensky’s original analysis.

<b>Input:</b> sila <sub>cl</sub> naglútò 3P.NOM AV.BEG-cook	NON-INITIAL (cl,IP)	LEFTMOST (cl, IP)
a. $\mathcal{C}_{IP}$ [nagluto sila <sub>cl</sub> ]		*
b. IP[sila <sub>cl</sub> nagluto]	*!	

**Tableau 2.3.** 2P clisis (Anderson 2000, 2005)

Central to most OT analyses of 2P clitics is Anderson’s idea that clitics represent the direct spell out of morphosyntactic features rather than terminal nodes in the syntax. In *A-Morphous Morphology* (Anderson 1992), clitics are features that are attached to phrase edges via the equivalent of word-formation rules. As a “late-insertion” model of morphology, these features are spelled out after regular syntax applies and their phonological forms can thus be sensitive to the presence of other features in the derivation. The OT account differs from Anderson’s earlier treatments in abandoning derivational rules in favor of a parallel evaluation. This means that, rather than being added sequentially by word formation type rules, all relevant features are co-present within the morphosyntactic representation. In this way, it is possible to account for the cooccurrence restrictions typical of clitic systems cross-linguistically. Furthermore, the internal ordering of multiple clitics is handled by extra-syntactic mechanisms which resemble word formation more than they do syntactic concatenation (Perlmutter 1971). For instance, person features and phonological weight are only emergent factors (if present at all) in the arrangement of bona fide syntactic material, but are able to constitute primary principles in the realm of clitic ordering.

On this view of things, many functional categories which are standardly taken to head phrases in GB/Minimalism are seen to be no more than features associated with clauses. This implies that such elements have a distinctly “unsyntactic” distribution and cannot take part in regular syntactic processes. In support of this,

Legendre (1998, 2000, 2001) argues at length that clitic auxiliaries in several Balkan languages differ from lexical auxiliaries in being syntactically inert. For instance, Bulgarian clitic auxiliaries, unlike lexical auxiliaries, do not partake of Subj-Aux inversion, do not count for the Head Movement Constraint, cannot be separated from the past participle by adverbs or floating quantifiers, and cannot be moved independently. A dividing line is thus posited between syntactic elements and purely morphological ones.

This approach also does away with the need to link clitic ordering to the position of corresponding phrases in the syntax. While the arrangement of phrases may be handled by standard operations such as MERGE and MOVE, the ordering of clitics is handled by an array of EDGEMOST constraints which target particular morphemes, again putting clitics on par with affixes (cf. Prince & Smolensky 1993). Thus a ranking in which LEFTMOST (cl<sup>1</sup>, IP) dominates LEFTMOST (cl<sup>2</sup>, IP) will result in the order [cl<sup>1</sup> cl<sup>2</sup> IP] where the numerals refer to morphological features of the clitic. This is demonstrated in Tableau 2.4 below, where we assume the input consists of fully formed words for ease of exposition.

<b>Input:</b> nila <sub>cl</sub> kami <sub>cl</sub> b<in>áti-Ø 3P.GEN 1P.NOM <BEG>greet-PV	NON-INITIAL (cl,IP)	LEFTMOST (cl <sub>GEN</sub> , IP)	LEFTMOST (cl <sub>NOM</sub> , IP)
a. $\leftarrow$ IP[b<in>áti-Ø nila <sub>cl</sub> kami <sub>cl</sub> ] 'They greeted us'		*	**
b. IP[b<in>áti-Ø kami <sub>cl</sub> nila <sub>cl</sub> ]		**!	*
b. IP[nila <sub>cl</sub> kami <sub>cl</sub> b<in>áti-Ø]	*!*		*

**Tableau 2.4.** Cluster-internal ordering of clitics

In (a), the optimal candidate, both clitics follow the first word in IP with the genitive clitic preceding the nominative one. This violates LEFTMOST constraint once for the genitive clitic and twice for the nominative clitic because they are separated from the left edge of IP by one and two morphological words, respectively.

Nonetheless, this solution is superior to switching the order of clitics as in (b), as this would violate the higher ranked LEFTMOST constraint on genitive clitics more than it would the lower ranked one on nominative clitics. Placing the clitics directly at the edge of IP, as in (c), would better satisfy both LEFTMOST constraints but only at the cost of placing the clitics initially in their domain and thus violating the highest ranked constraint in the tableau, NON-INITIAL (cl,IP).

Although this analysis derives the basic facts in an elegant manner and can handle ordering patterns in clitic syntax which cannot be derived from regular syntax, there are also some problems in completely detaching clitics from the syntax as suggested by Anderson (1992 *et seq*). Specifically, the following clitic-syntax interactions pose difficult challenges for an account which treats 2P clitics as “pure morphology”:

- (i) **Lack of clitic doubling** – clitics must be able to saturate syntactic requirements for overt arguments and to disallow cooccurrence with full DPs.
- (ii) **Inclusive construction** – nominative case clitics govern genitive case on an inclusive DP.
- (iii) **Coordination and modification** – 2P clitics may be conjoined with full DPs and modified under certain circumstances
- (iv) **Coercion** – 2P clitics may coerce lexical modifiers and DPs into 2P
- (v) **Impenetrability phenomena** – certain syntactic configurations cannot be interrupted by 2P clitics.

If the saturation of argument structure must be mediated by syntax, i.e. by the filling of syntactic positions by phrasal material, then, on a purely morphological

account of clitics, we would expect agreement-like behavior rather than complementary distribution between clitics and phrasal arguments.<sup>49</sup> Perhaps more difficult yet is the case relation seen in the inclusive construction and the apparent possibility for conjoining clitics with DPs. Both of these phenomena are challenging for “naïve” late-insertion models. In the first instance, the case of a DP should be determined in syntax while clitic case should be a late phenomenon. Yet, in the inclusive construction we find that it is the clitic which takes the expected nominative case and the inclusive DP whose genitive case is governed by the clitic. In the second instance, late-insertion models predict that clitics would not yet be visible at the time coordinated phrase are being built in the syntax and thus clitic-DP coordination should be impossible. Even more surprising for the morphological approach is the fact that 2P clitics can coerce DPs and modifiers into 2P, a position which was demonstrated above to not be accessible via syntactic movement. Finally there is the problem of impenetrability, to which an entire chapter is devoted to here. If clitics are inserted late in the derivation, somewhere in PF, we do not expect them to be sensitive to hierarchical syntactic structure. Yet, there exist certain syntactic configurations which are impenetrable to clitics and we find that there must be an overt connection in surface structure between the position of a pronominal clitic and the predicate whose argument structure it saturates. This is again not only unexpected but perhaps impossible to account for on a theory where clitics are completely divorced from syntax.

---

<sup>49</sup> There are of course many well known proposals for how dependent morphological elements (affixes and clitics) can satisfy argument structure requirements. These typically involve incorporation of a pronoun from argument position into the verb (Baker 1988, Bresnan & Mchombo 1986, 1987, Jelinek & Demers 1994, Anderson 1982) and thus require merging of the pronoun in a theta position. Anderson (2005), in part, follows Baker (1995) in positing a null category in argument position for so-called “pro-drop languages”. He differs from Baker in deriving the lack of (clitic and agreement) doubling by referentiality of the pronominal morphology rather than case absorption. In languages where pronominal morphology is referential, Anderson (2005:228-335) argues that cooccurrence with a full NP argument incurs a condition C violation. Based on facts from the inclusive construction, I will argue that the case based approach makes the most sense for Tagalog.



In the following section, the requirements of a descriptively adequate theory are brought together and the framework which will be adopted here is outlined. While this framework is still optimality theoretic in essence and involves elements from late-insertion models, it also permits broader interactions between inflectional morphology and syntax which are necessary to account for the problematic phenomena listed above.

## **2.4 An improved approach: 2P clitics as the output of Merge [F]**

### *2.4.1 The morphological framework*

Tagalog 2P clitics must be visible to certain syntactic processes such as coordination while making reference to surface phonological facts, such as syllable count (see Appendix), for ordering within the clitic cluster. Anderson's OT account, as other late-insertion accounts, appears to rule out the necessary access to syntactic structure (by virtue of current manifestations of the Bracket Erasure Convention (Kiparsky 1982)).

Here, both 2P clitics and more standard lexical items are merged in the syntax with the only major difference being that 2P clitics are purely featural elements which are adjoined to phrases while standard lexical items are more contentful and can only be merged as syntactic terminal nodes. The operation Merge can thus be understood to apply minimally to morphosyntactic features and maximally to full lexical items.<sup>50</sup>

Lexical items are merged to create syntactic structure in the familiar way, but it is less obvious how and where morphosyntactic features are merged. Here, I propose that features may be merged in two ways: (i) they may be merged to the functional

---

<sup>50</sup> Note that Merge [F] parallels Chomsky's (1995) proposal by which the operation Move may apply to features (Move-[F]) as well as to syntactic constituents. Perhaps due to lexicalist assumptions, the extension of feature manipulation to Merge as well as Move is dismissed by Chomsky (1995:262) as "vacuous" without further discussion.

phrases to which they relate and in this way may check features of functional categories, or (ii) they may be merged within a terminal node as in standard accounts (i.e. as in Hale & Marantz 1993 and subsequent work in Distributed Morphology). Under the first scenario, bundles of pronominal argument features can be merged directly to the edge of TP, where they are able to check T's case requirements.<sup>51</sup> The checking of case on TP by direct merger of clitics accounts for the case distribution facts in the inclusive construction. Because satisfying the case requirements of TP directly via (external) Merge of the relevant feature bundle is more economical than moving a DP for case checking purposes (cf. Chomsky 1995, 2000, McCloskey 2002 on the relative economy of Merge and Move), clitics have priority for case checking over full DP arguments.<sup>52</sup> Recall from §2.2 that when a full DP subject co-occurs with a clitic in the inclusive construction it is the clitic which reflects the expected nominative case and the full DP which takes default genitive, as shown in (108)a. If phrases took precedence over clitics for case checking purposes we would expect the ungrammatical case distribution shown in (108)b.

---

<sup>51</sup> Note that this diverges from Chomsky's tentative conjecture concerning the possible targets of Move [F]:

“Suppose that the target K is nonminimal. A reasonable conjecture is that the object formed, with a feature adjoined to a pure (nonminimal) maximal projection, would be uninterpretable at LF; independently, we will see that there are empirical reasons to suppose that an element adjoined to nonminimal K is not in the checking domain of its head H(K), so that the operation would be barred by Last Resort.”  
(Chomsky 1995:271)

With the subsequent development of case checking via Agree and the possibility of probing upwards as well as downwards for agreement (Baker 2008), it is not clear that Chomsky's evidence against adjuncts being outside the case checking domain of a lower head carries over to case checking by feature adjunction, although this cannot be fully evaluated here.

<sup>52</sup> Although this does not include the possibility of case checking by Agree without Move (Chomsky 2000). I am not aware of any discussion in the literature on the comparative economy of External Merge versus Agree.

- (108) a.     K<um>a~kain=pa=sila           ni=Edwin  
           <AV>INCM~eat=STILL=3P.NOM P.GEN=Edwin  
           ‘They ⊃ Edwin are still eating.’
- b.     \*K<um>a~kain=pa=nila           si=Edwin  
           <AV>INCM~eat=STILL=3P.GEN P.NOM=Edwin

In the inclusive construction, the DP functions semantically as a type of modifier of the pronoun, specifying a proper subset of its denotation and is thus licensed semantically. Although the inclusive DP cannot be assigned nominative case, it does have the option of being assigned default genitive case, which is assigned to a wide range of modifiers in Tagalog. The inclusive DP can thus be considered as being semantically licensed, but not case-licensed, as such. However, when the full DP is completely co-referential with a pronoun, the DP cannot be considered to be in a modification relationship with the pronominal and is thus neither licensed semantically nor syntactically, as nominative case is checked by the clitic. Clitic doubling is thus ruled out in Tagalog, as shown in (109), regardless of the case of the full DP.

- (109) \*K<um>a~kain=pa=siya           si/ni=Edwin  
           <AV>INCM~eat=STILL=3S.NOM P.NOM/P.GEN=Edwin

We have been tacitly assuming that the case of both External and Internal arguments of transitive clauses are checked at TP but this is highly unlikely for ergative case, which has been widely argued to be an inherent case associated with a lower projection in the “theta-domain” (vP for Aldridge (2004) and others, nP for Kaufman (2009)). The fact that both nominative and genitive (ergative) clitics appear to be adjoined to TP is understood here as a consequence of the requirement clitics must generally cluster together in Tagalog and cannot appear in separate positions in



pronominal cliticizes to the verb directly (enclitically in realis and proclitically in irrealis).

- (110) Moma=mo=**ko**      naria i-kalio=**ku**      *Kulawi*  
NEG=ALRD=2S.NOM EXT PV-see=1S.GEN  
'I don't hear you any more.' (Adriani & Esser 1939:17)

Thus, the positioning of genitive clitics higher than expected in Tagalog and other Philippine languages is best seen as a result of the need for clitics to cluster together in a single position. We revisit this process below in the discussion of the cluster internal ordering.<sup>54</sup>

We return back now to our second scenario above, in which features are merged within terminal syntactic nodes. Following the general idea put forth by Everett (1996), the type of Spell-Out which a particular feature bundle receives is partly dependent on where the feature bundle is merged. Everett argues that the difference between affixal agreement, pronominal clitics and free pronouns partly boil down to whether phi-features are merged within a word (affixes), adjoined to a word (clitics) or as independent syntactic terminals (free pronouns). While I do not endorse here this potentially over-simplistic difference between agreement affixes and clitics, the distinction between free pronouns and clitics as one of direct feature adjunction versus feature merger within an independent terminal node is adopted here and will play a large role in the chapters to come.

We can now turn to the coercion facts seen earlier with modification and to a lesser extent coordination. As mentioned above, these facts are problematic to most universal categorial theories of clitics as well as for late insertion approaches, as the

---

<sup>54</sup> If the need to maintain a contiguous cluster is responsible for the higher than expected position of genitive clitics, we expect them to surface in a lower position when appearing alone, but this does not occur. I leave this problem to further investigation.

clitic should not yet be visible to the syntax at the point where constituents are being conjoined. On the current approach, coordination of features with bona-fide lexical material also presents a dilemma, although it is not so much a modularity problem, as it is in Anderson's theory, but rather a simpler problem of linearization.

Let us assume that the phi-features of arguments, including person, number and case features, may be freely adjoined as a bundle to any position where they can be case licensed. Case licensing of arguments can take place through two main strategies: (i) in argument position via Agree (Chomsky 2000), or (ii) by direct adjunction of a feature bundle to TP, as suggested above.

In §2.2.2 we noted an interesting generalization concerning clitics and conjunction, the relevant facts of which are repeated below in Table 2.5.


**Table 2.5.** Coordination facts

<b>Construction</b>	<b>2P</b>	<b>Arg. position</b>
(i) Coordination		
a. ?	=cl & DP	...
b. *		... =cl & DP
c. *	=cl	... & DP
d. ✓	=cl	& DP

Whereas pronominal clitics could not be conjoined with full DPs in argument position and were judged by speakers to be marginal in unambiguous 2P, they appeared to be completely acceptable in a linear position which was ambiguous between 2P and argument position. In other words, conjunction was licensed so long as both conjuncts could satisfy their independent linear requirements. A relevant example is given in (112), repeated from (13).

- (112) Kamusta=ka at si=Mikki?  
 how=2S.NOM CONJ P.NOM=M.  
 'How are you and Mikki?'

This provides interesting support for surface or “phonological” oriented approaches to the Coordinate Structure Constraint (Ross 1967:89) in the form of inverse evidence. Grosu (1981:56) and Merchant (2001:193-200) claim that movement of a conjunct out of a coordinated structure is constrained by a PF ban on null conjuncts. The Tagalog data above shows that, in the extremely rare case where movement of a conjunct *into* a coordinated structure is independently motivated at PF, an otherwise ungrammatical configuration is licensed, as shown schematically in (113).<sup>55</sup>

$$(113) \quad [F]_{-TP}[PWd \text{ DP}[\text{___} \& \text{DP}]] \rightarrow [F]_{-TP}[PWd \text{ DP}[=cl \& \text{DP}]]$$


In this manner we can begin to explain surprising morphology interactions between syntax and 2P clitics. More difficult to explain are the cases of clitic coordinated structures in unambiguous 2P which are considered to be less than perfect by speakers but at the same time are not difficult to find attestations of. Here we must rely on the process of optional clisis, as discussed by S&O (1972: 184) and Billings (2005) whereby proper names can be placed in clitic position under certain pragmatic conditions (e.g. givenness, for Billings 2005). The optional positioning of a proper noun in 2P is shown in (114)b.

(114)a. Hindi  $\emptyset$ -dá~ratiŋ búkas si=Juan.  
 NEG AV-INCM-arrive tomorrow P.NOM=Juan  
 ‘Juan won’t be coming tomorrow.’

b. Hindi si=Juan  $\emptyset$ -dá-ratiŋ búkas.  
 NEG P.NOM=Juan AV-INCM-arrive tomorrow  
 ‘Juan won’t be coming tomorrow.’

<sup>55</sup> Several important questions surrounding these facts must remain open here. The three most glaring issues are why adverbial clitics may follow the pronominal clitic within the clitic cluster (e.g. *kamusta=ka=na at si=Mikki?*) if surface adjacency is at issue; how nominative case can be licensed on the second conjunct given the process of nominative bleeding by feature adjunction in the inclusive construction; and, if syntactic adjacency is more generally a PF phenomenon why can clitic placement only have an additive effect, as in (113), and never a subtractive effect, i.e. by triggering ungrammaticality when intervening between elements which must be syntactically adjacent?

Optional clisis however cannot apply to non-proper DPs, but only to personal names preceded by one of the personal name case markers (*si, ni*, etc.). Thus, while it may help explain conjoined proper names in 2P as in (115)a it will not offer much help when the element in question is not a proper name as in the inclusive construction in (115)b.

(115)a. Hindi=**ako** **at** **si=Juan** Ø-dá~ratiŋ  
 NEG=1S.NOM CONJ P.NOM=J. AV-INCM~arrive  
 ‘Me and Juan won’t arrive’

b. Hindi=**kami** **naŋ=pamílya=ko** Ø-dá~ratiŋ  
 NEG=1P.NOM GEN=family=1S.GEN AV-INCM~arrive  
 ‘Me  $\supset$  my family won’t arrive’

An interesting hint comes from Lebeaux (1988, 1990) who proposes that adjuncts can be merged late in the derivation, thereby explaining why certain adjuncts appear to be immune to condition C effects.<sup>56</sup> Essentially, Lebeaux’s proposal seeks to merge the relative clause to the already displaced *wh*- phrase, thereby avoiding a configuration in which the R-expression ‘John’ is c-commanded by a co-referring pronoun.

(116) [Which claim [that John<sub>j</sub> made]]<sub>i</sub> did he<sub>j</sub> later deny *t<sub>i</sub>*?

What constrains late merger? Because it is assumed that new syntactic relationships cannot be formed after late merger it follows that there can be no case and theta interactions between a late merged fragment and the rest of the clause. In fact, this delimits quite well the types of phrasal constituents which are able to be coerced into 2P in Tagalog: elements which both (i) share a theta role with the clitic, either by

---

<sup>56</sup> I am grateful to John Whitman for bringing the possible relevance of Lebeaux (1988) to my attention.



conjunction, inclusion or relative-like modification and (ii) share case with a clitic or are assigned default case, as in the inclusive construction. It is thus only by satisfying the theta and case requirements parasitically that bona-fide lexical material can be late merged in a non-syntactic position. This explains why essentially the same elements which are targeted for late merge in Lebeaux's theory, notably relative clauses, are those which can appear in 2P in Tagalog.

- (117) hindi=**nátiŋ**      **maŋa=banyágà** na-rá~ramdam-an      aŋ=maŋa=pátáy-an  
 NEG=1+2P.GEN:LNK PL=foreigner    NVL.BEG-INCM~feel-LV    NOM=PL=kill-LV  
 'We foreigners don't feel the killings.'

This also explains why such elements may only occur when coerced by a clitic element and not on their own, as underscored by the pairs in (118) and (119). Bona-fide lexical material cannot enter into a feature checking relationship by adjunction in the same way that purely functional feature bundles are claimed to here.

- (118)a. Hindi=**kami** **naŋ=pamílya=ko** Ø-dá~ratiŋ  
 NEG=1P.NOM GEN=family=1S.GEN AV-INCM~arrive  
 'Me ⊃ my family won't arrive'
- b. Hindi [**\*aŋ=pamílya=ko**] Ø-dá~ratiŋ [**aŋ=pamílya=ko**]  
 NEG NOM=family=1S.GEN AV-INCM~arrive NOM=family=1S.GEN  
 'My family won't arrive.'
- (119)a. Hindi=**sila=ŋ**      **lima** Ø-dá~ratiŋ  
 NEG=3P.NOM=LNK five AV-INCM~arrive  
 'They five won't arrive.'
- b. Hindi [**\*lima**] Ø-dá~ratiŋ      aŋ=[**lima=ŋ**] estudyante  
 NEG five AV-INCM~arrive NOM=FIVE=LNK student  
 'The five students won't arrive'

We are now much closer to making sense of the data summarized in Table 2.2

earlier. Interestingly, the one truly questionable case of coercion was that of clitic coordination with a DP and this is the one instance which requires an extra mechanism for case transmission between the clitic and the full DP. It is possible that such DPs in 2P are unacceptable for many speakers precisely because they have no way of satisfying their case requirements. Note also that among the attested examples given in (14)-(18) above, only one, (14), contained a clearly case marked conjunct. The others were either proper names, which can undergo optional clisis or bare noun phrases unmarked for case.

One final question which requires comment here concerns typological variation: What determines the presence or absence in a given language of “special” clitics of the Tagalog type? The mechanism of Merge [F], is too basic an operation to parameterize as being present or absent in a given language. Furthermore, it can be implicated in word building and other phenomena not directly related to 2P clitics. Rather what is at stake here is the target of Merge [F] and the value of [ $\pm$ EPP] features on functional projections requiring certain positions be filled syntactically, an aspect of the theory which has been argued to be independently necessary (Chomsky 2000) (although see Grohmann 2000 et al for another view). In a language like English, Merge [F] only targets  $X^0$  nodes adding phi and case features to syntactic heads. In languages like Tagalog, these features are merged separately from the syntactic heads which they are ultimately associated with in English. Such a model of structure building is at odds with Chomsky’s (1995) lexicalist view of insertion but does not deviate significantly from the grammatical architecture espoused by Distributed Morphology. The main difference between Merge-[F] and the Distributed Morphology conception of word building (as well as that of Baker 1985) is that, in the latter theory, words are thought to contain full fledged X-bar structures, regardless of whether there exists evidence for a Specifier position, additional adjunction sites or other hallmarks

of the X-bar configuration. While not ruling out that possibility in principle for other pieces of morphology, argument features are considered here to be best treated not as originating in being associated with dedicated functional phrases but rather as being added directly from the numeration either to existing syntactic heads or phrase edges.<sup>57</sup>

This concludes the brief exposition of the morphosyntactic framework which will be relied on here. In the next subsection, an overview of the phonological framework is presented with particular attention given to why the phonology militates against spelling out clitics in first position.

#### 2.4.2 *The phonological framework*

The basic framework which will be adapted here for the phonology as well as for the spell out position of clitics is OT. As discussed above, OT accounts see 2P phenomena as the result of a conflict between a constraint which requires that a given element surfaces leftmost in its domain and another constraint which militates against absolute initial position. Here, we will argue that the LEFTMOST constraint is actually a constraint which aligns the spell out of morphemes with their underlying position. The underlying position of clitics is derived by Merge [F], as proposed earlier, and delineates the leftmost boundary of the range that a given clitic may appear in. As will be explained more thoroughly in chap. 4, ALIGN MORPH ensures that feature bundles

---

<sup>57</sup> This supposes that there is no direct connection between clitic position and clitic doubling which I take to be a typologically correct observation. Based on Slavic and Romance data it has been claimed that 2P clitics categorically disallow doubling (Halpern & Fontana 1994, Rivero 1994, 1997, Franks & King 2000) but this is clearly incorrect, as shown by the existence of doubled 2P clitics in Kapampangan and the South Sulawesi languages, among others (Billings & Kaufman 2004, Kaufman to appear). The typological tendency for 2P clitics to disallow clitic doubling is due to doubling and cohesion to a host both being subject to similar forces of grammaticalization; the closer a pronominal/agreement element is to its host the more chance it has of being ignored for theta role assignment. Nonetheless, the existence of numerous counterexamples to the 2P-doubling generalization demonstrates that it is orthogonal to UG.

are spelled out where they are merged. The overall vision of the grammar is thus one in which syntax builds a syntactic and morphological structure via Merge (Terminal) and Merge [F] which is later sent to Spell Out, a process mediated by violable ranked constraints as per OT. These constraints are primarily of a phonological nature but may also make reference syntactic structure in a limited way.

In the following subsection we concentrate on the phonological side of Spell Out examining the nature of non-initiality for Tagalog 2P clitics. It is argued that Anderson's NON-INITIAL constraint is best viewed as prosodically, rather than morphologically, motivated.

#### 2.4.3 *Deriving 2P*

Anderson notes that there is an apparently serious problem with treating 2P clisis as a phonological phenomenon in Tagalog:

“While there may be languages for which a phonological account of the non-occurrence of clitics in phrase-initial position is possible, no well supported analyses of this kind have been presented. And in fact it is extremely unlikely that such a prosodically based account will be adequate in general. That is because some special clitics that must be placed post-initially are not prosodically deficient. Tagalog, for instance, has a huge system of clitics, most of which are prosodically autonomous and bear their own stress. There seems no phonological reason why these could not occur initially, and if they do not, that fact must be due to some other constraint.” (Anderson 2005:141)

Anderson thus suggests that non-initiality be derived with a *morphological* constraint NON-INITIAL. Although he does not provide explicit evaluations, the one given in Tableau 2. earlier sums up his analysis. Although NON-INITIAL makes correct predictions for the basic cases, it also represents a stipulation, as NON-INITIALITY is not made to follow either from more general phonological or syntactic principles. As

Anderson himself notes, we also erroneously predict that non-initial be mirrored by a NON-FINAL constraint:

“Whatever the role of **NonInitial (e)** and **LeftEdgeFaith (P)** in the description of morphological and phrasal affixes, their mirror images **Non-Final (e)** and **RightEdgeFaith (P)** are much less prominent, if indeed they exist at all. It is possible to suggest as a motivation for this asymmetry that the identification (and stability) of left edges is important in itself for morphosyntactic parsing, while no corresponding significance is attached to right edges.” (Anderson 2005:142)

While some evidence in support of a morphological NON-FINAL has been discussed in the literature (Sityar 1991, Franks & King 2001, Anderson 2005:150), the data appears inconclusive and at most it is only an emergent tendency which plays no part in the typology of clitic placement, as there are no penultimate clitics (2P from the right edge) (see chapter 3). Thus, just as Anderson claims that there is no solid universal prosodic account of initial position avoidance, there is also no existing morphological account which would disallow reference to right edges. The burden is simply shifted to the morphology in an account which relies on NON-INITIAL. The problem which Anderson notes, however, is quite real and must be dealt with by any prosodic account for initial position avoidance. If 2P clitics cannot be identified as prosodically deficient, then there can be no prosodic basis for removing them from initial position. This deficiency can of course be stipulated in the lexicon as a Prosodic Subcategorization requirement ala Inkelas (1990). This was seen above in Chung’s account of Chamorro clitics and is also akin to the strategy opted for by Billings (2004), who employs the following constraint in his own OT analysis of 2P clisis:

SUFFIX  
Morphemes marked as suffixes must follow some PWd

Billings' account relies partly on the combination of a constraint enforcing intonation phrase alignment, across which suffixation cannot take place, and a lexical specification for 2P clitics as [+suffix]. Nonetheless, if there exist independent non-clitic homophones of 2P clitics and these clitics otherwise satisfy all requirements for prosodic wordhood, as in Tagalog, we are still left with a purely diacritic notion of clitic. Additionally, the nature of the host (the PWd) is built into the suffix constraint itself, as it is in Prosodic Subcategorization frames, but optimally should be derived by independent means.

As in the citation above, prosodic weakness is often equated with prosodic dependency, but, logically, this need not be the case. Returning to our infixation analogy, the vowel of Tagalog <um> should not be characterized as possessing an inherent dependence on a preceding consonant, but can rather be thought of as a “weak” element (in an abstract sense) which *preferably* follows a “strong” element. We can further recognize the universal tendency for prosodic constituents to start “strongly” (cf. Kavitskaya 2005, Smith 2002, Beckman 1998). Fortition on the syllable level has been observed to only occur in initial position and never in final or medial positions (Hooper 1976:199). On the level of the metric foot, we may cite the preference for trochees in acquisition (Fikkert 1994, Demuth 1995, Pater 1997) and the overall preference for trochaic over iambic meter cross-linguistically (Hyman 1977). On the level of the prosodic word, Kavitskaya (2005) discusses cases of regular prosodic word initial fortition, as in Kurdish (Bradley 2001, Abdulla & McCarus 1967) and we may also include here well known evidence for initial dactyl effects (Prince 1983, Hayes 1985, Cohn 1993, Cohn & McCarthy 1994)<sup>58</sup>, by which a prosodic word must be stressed on its initial syllable. Finally, subphonemic strengthening has also been found by Keating et al (2003) and Fougeron & Keating

---

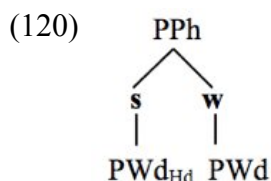
<sup>58</sup> I thank Draga Zec for bringing the relevance of initial dactyl effects to my attention.

(1997) to occur consistently on the onsets of prosodic and intonational phrases in a wide range of languages. Selkirk (1980) and Hayes (1981) have also argued that the prosodic hierarchy contains dyads containing weak and strong members throughout.

The universal tendency towards domain-initial fortition is formulated here as the outcome of \*WEAKSTART:

\*WEAKSTART (PCat)  
 Violated by the configuration [**w** (**w**)...] in PCat  
 (Do not begin a prosodic domain with a weak element)

This constraint assigns violations to prosodic categories whose left edges are aligned to weak elements and, in this sense, is perhaps best conceived of as a generalized ONSET constraint.<sup>59</sup> Applying this constraint template to the prosodic phrase we can derive (120) as the optimal configuration, in which a prosodic word head is aligned to the left edge of its containing prosodic phrase.



Heads and non-heads on the prosodic word level are extremely similar in Tagalog, thus giving rise to observation that they are not prosodically dependent as clitics are often described to be. There is one clear manifestation, however, of non-head status and that is the impossibility of bearing prosodic marked focus, as exemplified in (121).<sup>60</sup>

<sup>59</sup> See Ito & Mester 1999 for a similar formulation. Unlike ONSET however, \*WEAKSTART can be violated multiple times within a single domain.

<sup>60</sup> Although this may also be due to independent constraints requiring focused elements to surface in predicate position.

- (121) %Nag-lútò=SILA  
 AV.BEG-cook=3P.NOM  
 (only interpretable as ‘[They cooked]<sub>F</sub>’ not ‘[They]<sub>F</sub> cooked’)

Focusability thus differentiates clitic and free pronominals. Free pronouns in initial position, as in (122), can receive focal prominence (without focus projection), unlike their clitic counterparts in second position.

- (122) a. [SILA]<sub>F</sub> ay=nag-lútò                      b. [SILA]<sub>F</sub> ang=nag-lútò  
 3P.NOM TOP=AV.BEG-cook                      3P.NOM NOM=AV.BEG-cook  
 ‘THEY cooked’ (contrastive topic)              ‘THEY cooked’ (focus)

Zec (2005:87) argues on the basis of Serbo-Croatian that the assignment of PwD head status is correlated to lexical word status. This is enforced by the constraint below, to which we have added the second clause, thus making it symmetrical:

PHead=MWd<sub>Lex</sub>  
 A PwD with prosodic head status must be a MWd<sub>Lex</sub>.  
 A MWd<sub>Lex</sub> must be a PwD with prosodic head status.

This constraint differentiates lexical from functional categories, enforcing head status on the former but not the latter. Following this line of analysis we can take the free set of pronominals and adverbials to be lexical words – understood here to mean those items occupying a (phrasal) position in the syntactic tree – as opposed to their weak counterparts, which are strictly composed of adjoined functional features.<sup>61</sup> Having sufficiently formalized the unmarked status of the s-w type for our purposes we can now focus on how \*WEAKSTART operates in prosodic phrases to yield 2P clitics. Assuming prosodic phrases are aligned to syntactic phrases (Truckenbrodt 1999), we can easily imagine how \*WEAKSTART (PPh) assigns violations to structures

---

<sup>61</sup> Supporting evidence for the distinction between true lexical items and the morphemes which result from the spelling-out of features is found in feature cooccurrence constraints, which only make reference to featural material and ignore lexical items (see chap. 4).



in which prosodic phrases begin with prosodic words which are not heads. This is shown in the trivial evaluation in Tableau 2.5, below.

<b>Input:</b> MWd <sub>Lex</sub> <i>naglútò</i> AV.BEG-cook  MWd <i>sila</i> 3P.NOM	(a) $\uparrow$ PPh 	(b) PPh 	(c) PPh 
	PHead=MWd <sub>Lex</sub>		
*WEAKSTART (PPh)		*!	

**Tableau 2.5.** The prosodic basis for non-initiality

With the above constraints, the feature which ultimately results in morphemes projecting prosodic word heads or not is [ $\pm$ Lex(ical)]. This, I claim can further be derived from the way that a particular item is merged into the syntax, as seen with morphemes which have homophonous functional and lexical counterparts, i.e. the Tagalog nominative pronouns. When merged as a terminal node, in particular, as complement to a head, an item is [+lexical]. On the other hand, an item merged as an adjoined feature bundle can only be [-lexical]. The strategy taken here has the important advantage of being able to place clitics in second position without requiring them to have inherent prosodic or morphological dependencies. It is rather the requirements of the prosodic phrase rather than the requirements of clitics which forces them into second position. Disyllabic 2P clitics can thus be properly characterized as “free Wackernagel elements” (cf. Nichols & Bickel 2007), or in the terms of Wackernagel (1892) himself, “Quasi-Enklitika”, i.e., items that were

“optionally enclitic but still occurred in Second Position” (cf. Zwicky’s (1977) and Nevis’ (1988) “bound word” category).

#### 2.4.4 *Cluster internal ordering*

We have been hitherto primarily concerned with the syntactic properties of clitics as they relate to positioning within the clause. The other crucial aspect of clitic morphosyntax which makes clitics “special” is how they are ordered in relation to themselves, also known as “cluster internal ordering”. Although we do not concentrate here on cluster internal ordering, no discussion of Tagalog clitics is complete without at least a brief exposition of the facts. Tagalog cluster internal ordering has drawn some attention because of its typological rarity (S&O, Schachter 1973, Billings & Konopasky 2002, 2003, Billings & Kaufman 2004), in particular, the strong role of syllable count in ordering clitics. If syntax were responsible for clitic ordering we would expect clitics to be linearized according to the order of their corresponding functional projections, either reflecting the base order, or the reverse order, as a result of successive left adjunction. In fact, neither of these predictions turns out to be correct. The relatively uncontroversial order of projections in (123)a indicate the clitic categories with which they should be associated on the basis of their function and meaning. High adverbs, such as the question marker and speaker oriented adverbs should be associated with the CP layer. Nominative case clitics should be associated with TP. Aspectual clitics should be associated with an aspectual projection between TP and vP, and (inherent) genitive case clitics should be associated with the lowest layer where arguments are base generated. As shown by (123)b, and c, neither the base order nor the reverse order are attested in Tagalog. Rather it is the phonologically based order in (123)d which is correct.

(123)a. CP[HighAdv TP[NOM AspP[ALRD vP/nP[GEN]]]]

b. \*ba=ka=na=niya  
QM=2S.NOM=ALRD=3S.GEN

c. \*niya=na=ka=ba  
3S.GEN=ALRD=2S.NOM=QM

d. ka=na=ba=niya  
2S.NOM=ALRD=QM=3S.GEN

Tagalog clitics follow a light-first ordering in which all monosyllabic clitics precede disyllabic clitics. Only if two pronominal clitics have the same syllable count – which, due to suppletion with the combination of *ko* 1S.GEN and *ka* 2S.NOM, only occurs with disyllabic clitics (but see one exception below) – then a genitive clitic must precede a nominative clitic. Thus, while person and case features determine ordering in the vast majority of attested clitic systems, Tagalog displays a clearly phonological ordering with case playing an apparently secondary role. As we have seen, adverbial clitics also form part of the cluster. These are ordered in relation to each other quite closely to what would be expected on a scopal basis (see Kaufman 2005, for the role of scope in the positioning of adverbials as well as clitics), with inner adverbials (e.g. aspectual modifiers) at the left edge of the cluster and outer adverbials (e.g. the question marker) on the right edge, as shown in (124). Note that the scopal ordering holds within each phonological domain but not across them. For instance, the question marker, which should be on the outer edge according to most theories of scope and adverbial projections, must still precede the emphatic/switch topic marker *naman*, because the latter is disyllabic and thus belongs to the disyllabic phonological domain.

(124) na, pa > na-naman > man > nà > din ~ lan > daw > pò > ba >  
ALRDY STILL AGAIN EVEN EMPH ALSO ONLY RPRT POLITE QM

naman > yàtà, pala, sàna > kayà  
SWITCH.TOPIC PERHAPS MIRATIVE OPTATIVE SPECULATIVE

It thus seems that the “reverse” linearization derivable by successive leftwards adjunction is closer to the attested order than the base order for both disyllabic pronominals of different cases as well as for adverbials. It could thus be assumed that leftwards adjunction plus a phonological filter could derive the correct linearization, but in fact, there exists an insoluble problem for such an approach. The greatest puzzle of Tagalog cluster internal ordering is the relationship between adverbial and pronominal clitics. While we might expect adverbials and pronominals to form distinct sub-clusters, we find that the adverbial sub-cluster is sandwiched by the monosyllabic pronominals on the left edge and the disyllabic pronominals on the right edge. The macro-ordering is shown schematically in (125).

(125)  $1\sigma \text{ PRON} > 1\sigma \text{ ADV} > 2\sigma \text{ ADV} > 2\sigma \text{ PRON}$

Although the template in (125) obeys several principles which are found to hold more generally, it proves impossible to combine these independent principles in a simple way to predict the correct pattern. The principles in question are the following:

LIGHT FIRST – monosyllabic clitics precede disyllabic clitics

GENITIVE FIRST – genitive clitics precede nominative clitics

SCOPE – internal adverbial clitics precede external adverbial clitics

Note that LIGHT FIRST is also seen to be operative in the phrasal syntax of the language but only as a tendency rather than a rule (Kroeger 1993, Billings 2005) and GENITIVE FIRST is the linear ordering which emerges from the constituency of Philippine predicate initial transitive clauses, where the genitive/ergative argument forms a phrase with the predicate head that excludes the nominative, i.e.,  $[[\text{PRED}^0 \text{ GENP}] \text{ NOMP}]$  (see Kroeger 1993, Kaufman 2009).

What is responsible for the clustering together of adverbials and pronominals? It cannot be narrow syntax since the clustering of pronominals on either side of the adverbials is dictated by their syllable count. Genitive and nominative clitics appear both on the left and right of the adverbials in order to satisfy the LIGHT FIRST constraint. Let us assume that the constraints ALIGN-L [PRON] and ALIGN-L [ADV] are responsible for the clustering together of adverbials and pronominals, respectively. We can see in Table 2.6 how a simple ranking of one of these constraints over the other with both being dominated by LIGHT FIRST, as is necessary, yields incorrect linearizations.

**Table 2.6.** Possible rankings for relative position of adverbial and pronominal clitics

RANKING	LINEARIZATION
a. LIGHT >> ALIGN-L >> ALIGN-L FIRST [PRON] [ADV]	*1σ PRON > 1σ ADV > 2σ PRON > 2σ ADV
b. LIGHT >> ALIGN-L >> ALIGN-L FIRST [ADV] [PRON]	*1σ ADV > 1σ PRON > 2σ ADV > 2σ PRON

Neither simple ranking can derive the fact that pronominals sandwich adverbials, as in (125). We could further assume that there exists a need for adverbial clitics to be contiguous but this cannot predict that the same orders emerge when only one adverbials is present, as shown by the orders in (126).

- (126)a. =ko=na                      b. =sana=kami  
1S.GEN=ALRD                      OPT=1P.NOM

For an OT account to predict the facts it appears necessary to employ the conjoined constraint (Smolensky 1993, Moreton & Smolensky 2002) LIGHT FIRST & ALIGN-L [PRON]. This is violated when a single element independently violates both LIGHT FIRST and ALIGN-L [PRON]. Constraint conjunction captures cases in which the violation of a certain combination of constraints is worse than would be predicted on

independent grounds, in other words, cases where the whole of a violation is larger than the sum of its parts.<sup>62</sup> To see how the conjoined constraint can yield the correct outputs for two basic cases, observe the evaluations in Tableau 2.6 (the disyllabic case) and Tableau 2.7 (the monosyllabic case).

Input: OPT, 1P.NOM	LIGHT FIRST	LIGHT FIRST & ALIGN-L [PRON]	ALIGN-L ADV	ALIGN-L PRON
a. <i>kami sana</i> 1P.NOM OPT			*!	
b. ☞ <i>sana kami</i> OPT 1P.NOM				*

**Tableau 2.6.** Cluster-internal ordering, simple disyllabic case

Input: 1S.GEN, ALRD	LIGHT FIRST	LIGHT FIRST & ALIGN-L [PRON]	ALIGN-L ADV	ALIGN-L PRON
a. ☞ <i>ko na</i> 1S.GEN ALRD	*		*	
b. <i>na ko</i> ALRD 1S.GEN	*	*!		*

**Tableau 2.7.** Cluster-internal ordering, simple monosyllabic case

The essence of this analysis is that adverbs naturally precede pronominals in Tagalog but that extra penalties are incurred when a pronominal is both light and misaligned from the left edge of the clitic cluster. In the case of two disyllabic clitics in Tableau 2.6, there are no light clitics and thus the conjoined constraint is inactive. However, in the case of two monosyllabic pronouns shown in Tableau 2.7, the LIGHT FIRST constraint is violated by the rightmost clitic. When that clitic is a pronominal it violates both LIGHT FIRST and ALIGN-L [PRON] and thereby also incurs a violation of the conjoined constraint LIGHT FIRST & ALIGN-L [PRON], which renders it worse than the candidate which misaligns the adverbial. When we add a disyllabic pronominal

<sup>62</sup> Constraint conjunction has engendered much discussion regarding computability and restrictiveness in OT. See Potts & Pullum (2002) for a summary and argument that certain types of conjunctions are computable.

and adverb, the “natural” adverb first ordering will prevail, as shown in Tableau 2.8, as the conjoined constraint only comes into play with monosyllabic clitics.

Input: 1S.GEN, ALRD, OPT, 3S.NOM	LIGHT FIRST	LIGHT FIRST & ALIGN-L [PRON]	ALIGN-L ADV	ALIGN-L PRON
a. <i>ko na sana siya</i> 1S.GEN ALRD OPT 3S.NOM	*		* na ** sana	*** siya
b. <i>ko na siya sana</i> 1S.GEN ALRD 3S.NOM OPT	*		* na ***! sana	** siya
c. <i>na ko sana siya</i> ALRD 1S.GEN OPT 3S.NOM	*	*!	** sana	*ko ***siya

**Tableau 2.8.** Cluster-internal ordering with mixed clitics

It is worth emphasizing that the phonological, weight based nature of clitic linearization in Tagalog is conclusively demonstrated by the placement of the suppletive clitic *kita* 1S.GEN+2S.NOM, as first pointed out by Schachter (1973). When not appearing within the same cluster, 1S.GEN and the 2S.NOM are positioned at the left edge of the cluster due to their being monosyllabic, as shown in (127)a and b. However, when appearing together, their suppletive counterpart *kita* appears at the right edge of the cluster as it is disyllabic, as seen in (127)c. The positioning of clitics must thus take place on the surface with access to the syllable count of suppletive allomorphs.

- (127) a. =**ko**=na=sana=sila  
=1S.GEN=ALRD=OPT=3P.NOM
- b. =**ka**=na=sana=nila  
=1S.NOM=ALRD=OPT=3P.GEN
- c. =na=sana=**kita**  
=ALRD=OPT=1S.GEN+2S.NOM

The cluster internal facts thus support a view of clitic positioning in which clitics can be simultaneously subject to both syntactic and phonological constraints. While the general syntax may play a role in the emergent genitive before nominative ordering, the phonology must be able to play the dominant role by sorting clitics

according to syllable count. Thinking derivationally, it is only after phonology makes the first cut with LIGHT FIRST that the morphosyntactic GENITIVE FIRST plays a role. Furthermore, “long distance” reordering of clitics must be able to see syllable count of suppletive morphs, contra late insertion models. As reviewed in §2.3.4, in Distributed Morphology only the processes of Local Dislocation and (perhaps) Prosodic Inversion can take place after lexical insertion. It is not exactly clear what would count as local within the context of the clitic cluster but it seems to be a safe assumption that this would be restricted to metathesis of adjacent items (see §2.3.4 for details), and thus be too restrictive. The disyllabic structure of the suppletive *kita* should therefore not be able to determine its position within the clitic cluster within such a conception of the interface. Here, both phonological and morphosyntactic features can freely enter the computation of how particular morphemes are linearized.

## **2.5 Conclusion**

In this chapter, we have reviewed numerous theories of clitic positioning and evaluated them on the basis of Tagalog facts. It was shown that Tagalog clitics have a curious *mélange* of properties which are exceedingly difficult to capture by any existing theory. Specifically, these clitics appear in positions which syntactic movement cannot access, similar to what has been argued by Chung (2003) for Chamorro 2P clitics. Regular syntactic mechanisms of clitic placement can thus be largely ruled out for Tagalog just as they are for Chamorro. On the other hand, impenetrability phenomena suggest the continued relevance of syntactic structure, a state of affairs which should be impossible if clitics are by the phonological process of Prosodic Subcategorization. Tagalog clitics were also seen to have surprising interactions with phrasal material, a fact which can only be explained by their participation in syntactic structure. The general stance developed in this chapter is that



2P clitics are the spell out of feature bundles merged in irregular positions by regular syntax (i.e. the operation Merge [F]). It will be shown in chap. 4 that the irregular status of adjoined feature bundles without a syntactic node makes clitics more susceptible to phonological pressures than other items. The precise spell out position of pronominal clitics will then depend on prosodic and syntactic factors. Specifically, these clitics must satisfy \*WEAKSTART while at the same time maintaining a visibility relation with the predicate head from which they receive their theta-role, as will be discussed in chap. 5.

## CHAPTER 3: THE TYPOLOGY OF CLITIC PLACEMENT

### 3.1 Introduction

Various parameters have been offered in the literature to determine the domain, position, and phonological parsing of clitics although very few works have attempted a thorough typological survey of the facts. Most famously, Klavans (1980, 1985, 1994) posited three binary parameters to determine clitic position and parsing: Initial/Final within a positioning domain, Before/After the pivot, and Proclitic/Enclitic attachment to the host. Crucially, these three parameters were argued to be completely independent of each other, thus yielding eight combinatorial possibilities. In this chapter, we use the Klavans typology as a basis for our own empirical investigations. Through a reevaluation of later work and examination of several cases which have not been considered, we are able to simplify the overall typology by reference to two morphosyntactic types of clitics: those which are in a syntactic head complement relationship (syntactic heads) and those which are not (adjoined feature bundles) (see also Werle 2004, 2009). This sets the stage for the dual mechanism approach to clitic syntax introduced in the following chapter whereby head clitics are understood as ordinary lexical items merged from the numeration while non-sister clitics are added by Merge [F] as feature bundles directly to existing phrase structure.

#### 3.1.1 *The Klavans typology*

To understand the behavior of these eight types more precisely, observe the schema in (1). The bracketed constituents in (1) represent the domain with which the clitic is associated. The letters A-D represent the clitics' morphological pivot with the equals sign indicating prosodic attachment on some level. (1)a illustrates the positions

within a phrasal domain and (1)b illustrates the positions when the domain is a head.<sup>1</sup> The possibilities are reduced when the domain is a head because of the (near) absence of endocclisis; lexemes are generally opaque for the purposes of clitic placement and thus the domain internal positions (i.e., positions 3-6) are absent here.

(1) a. **A=1** <sub>XP</sub>[**2=B=3** 4=...=5 6=**C=7**] **D=8**

b. **A=1** <sub>X<sup>0</sup></sub>[**2=B** **C=7**] **D=8**

These positions can be exemplified more concretely in (2) with the object pronoun *her* in the sentence, “Then, Jane drove **her** to work. Later...”. Whereas in English, the object pronoun has much the same (although not identical) distribution of a phrasal argument, Klavans claims that all the positions illustrated schematically in (2) are attested for clitics with a phrasal domain. All attachment sites are domain internal except for 1 and 8, which attach rightwards to the end of previous material and leftwards to the beginning of following material, respectively.

(2) Then=**her**<sub>1</sub>, <sub>TP</sub>[**her**<sub>2</sub>=Jane=**her**<sub>3</sub> **her**<sub>4</sub>=drove to **her**<sub>6</sub>=work=**her**<sub>7</sub>] **her**<sub>8</sub>=Later...

Klavans exemplifies the eight possibilities with the clitics listed in

**Table 3.1.**

---

<sup>1</sup> As discussed below, the difference between phrasal clisis and head-adjacent clisis is not distinguished clearly by Klavans. Her typology consequently suffers from this oversight and can be reduced significantly.

**Table 3.1.** Klavans (1980, 1985) 8-way typology

<b>P1 Initial/Final</b>	<b>P2 Before/After</b>	<b>P3 Proclitic/Enclitic</b>	<b>Example</b>
1. Initial (under N')	Before	Enclitic	Kwakwala NP markers
2. Initial (under N')	Before	Proclitic	Greek article
3. Initial (under S)	After	Enclitic	Ngiyambaa enclitics
4. Initial (under S)	After	Proclitic	Tepecano <i>an=</i>
5. Final (under S)	Before	Enclitic	Nganhcara clitics
6. Final (under S)	Before	Proclitic	Sanskrit pre-verbs
7. Final (under V[-T])	After	Enclitic	Spanish pron clitics
8. Final (under S)	After	Proclitic	Greek negative <i>ou=</i>

The restrictiveness of the Klavans typology stems from the restriction on the “pivot”, i.e. P1, which is limited to seeing only the edges of host constituents. Later studies, however, revealed several empirical problems with

Table 3.1 (Sadock 1991, Marantz 1993, Halpern 1995, Billings 2004) and consequently forced a revision of the typology. We review here in detail the evidence for each type of clitic and the problems encountered by Klavans’ parametric approach. Crucial in reducing the typology is the recognition of two independent distinctions in clitic type: head-adjacent versus phrasal (Billings 2004, Halpern 1995) and morphosyntactic sister versus non-sister. Sister clitics are often, but not always, functional heads selecting for lexical complements. Pronominals can also be sister clitics in their base position within *vP* but not when attached outside of *vP*. Non-sister clitics are primarily made up of pronominals outside of *vP* and adverbials (modals, aspectual markers, question markers, etc.).<sup>2,3</sup> Some typical exemplars of sister and non-sister clitics which will be featured in the discussion to follow are listed in (3).

<sup>2</sup> Auxiliaries occupy an intermediate position here for reasons which cannot be fully explored here. In respect to the typology being presently developed, they behave like sister clitics in some languages (e.g., English) and non-sisters in others (e.g., Serbo-Croatian). Because of this ambiguity we restrict our investigation to the clearer types.

<sup>3</sup> Certain “adverbial” elements can also be base generated functional heads, and therefore sister clitics. Question markers, for instance, can be generated as heads of a dedicated projection in the C-layer (e.g. Rizzi’s ForceP). As such, these need not be clitics at all and can in fact host 2P clitics in such languages as Maranao and Seediq (see (8) and (9) below). The diagnostic for sisterhood as understood here is a correspondence between surface constituency and semantic compositionality. Functional elements which, in their canonical position, are dislocated from the constituent with which they compose

- (3) **Sister clitics:** case markers, adpositions, Comp, pronominals within  $vP$   
**Non-sister clitics:** adverbials, definiteness markers, pronominals outside  $vP$

The rest of this chapter is organized according to Klavans' typology. Because it is already clear that head-adjacent clisis can be either to the left or right edge of the host, we concentrate here on phrasal clisis. In order to test the simple-special dichotomy we compare unambiguous phrasal heads, such as case markers and adpositions with pronominals, i.e. non-phrasal heads. We affirm some of the conclusions reached by earlier critics of this typology and further incorporate the notion of markedness into the resultant typology based on the apparent extreme rarity certain types. Ultimately, this will allow our analysis to make predictions not only about possible and impossible clitic types but also to model the reasons behind the cross-linguistic rarity of other attested types. The following sections take a closer look at each clitic type while bringing in new data primarily from Austronesian languages.

## 3.2 Clitic types

### 3.2.1 *Type 1: initial, before, enclitic*

The phenomenon of enclitic prepositions was famously discussed by Sapir in his grammar of K<sup>w</sup>ak<sup>w</sup>ala and has served since then to demonstrate the “independence of phonology from syntax” (Klavans 1985). Billings, Kaufman & Werle (2006) employ Limos Kalinga (North Cordilleran; N. Philippines) to exemplify this type, shown in (4) and (5). The relevant element here is the *prepositional* oblique case marker which has two allomorphs /si/ and /t/ the choice of which is conditioned by *preceding* phonological material. If the preceding material ends in a vowel the case

---

semantically, are understood to be non-sister clitics. Clitics which are in their “expected” position are sister clitics. Note that this differs from Zwicky's simple/special distinction in that, here, pronominal sister clitics do not necessarily have the syntax of their NP counterparts but need only be sister to the verb (or other host within the verb complex).

marker surfaces as /t/ whereas it surfaces as /si/ if the preceding material ends in a consonant. In this case, which is quite common among languages of the Northern Philippines, the case marker cliticizes to the left despite being syntactically related to the following material.<sup>4</sup>

- (4) Nan-anup dadit tagu=t bolok *Limos Kalinga*  
 AV.BEG-hunt PL.NOM person=OBL pig  
 ‘The people hunted pig.’ (Ferreirinho 1993:12)
- (5) Mam-mula=ak si balat. *Limos Kalinga*  
 AV-plant=1S.NOM OBL banana  
 ‘I’m planting bananas.’ (Ferreirinho 1993:82)

The Kalinga case marker is a syntactic head which appears in the expected position, i.e. as a simple clitic in Zwicky’s terms. Cases of similar “encliticizing prepositions/preposed case markers” abound in Philippine languages and are also widely outside the Philippines.

What has never been clearly shown to exist are non-sister clitics with the same positioning properties. The closest we come to such a case is the Warlpiri present tense clitic *ka*, which normally appears in second position, as shown in (6).

- (6) Njuntu=**ka**=rna=ngku kuyu-ku yilya-mi *Warlpiri*  
 you=PRES=1.SUBJ=2.OBJ meat-JUSSIVE send-NONPAST  
 ‘I am sending you for meat.’ (Anderson 1993:82)

However, Simpson (1991:69) (via Billings 2004) states, “in connected speech, monosyllabic AUX bases are found sentence initially, because the last element of the previous sentence provides a phonological host for the clitics.” One of the very few

---

<sup>4</sup> Billings (2004) exemplifies this type with the Tagalog linker, which can be shown to be a prepositional head whose allomorphy /na/ vs. /ŋ/ also depends on preceding material.

examples of this type of attachment is cited from Hale (1966) (p.c. Jane Simpson to Stephen Anderson). The example in (7) is cited from Legate (2008:10).

- (7) Jinarn-kiji-ni-ji.                      Ngula-nya-ka-rnalu                      ngarri-ni  
trip-throw-NONPAST-TOP    that-FOC-PRESIMPF-1P.EX.SUBJ    call-NONPAST
- jinarn-kiji-ni-ji,                      kaji-lpa-npa  
trip-throw-NONPAST-TOP    NONFACT-PASTIMPF-2S.SUBJ
- watiya-rla-rlangu    wanti-yarla. **Ka**    jinarn-kiji-ni.  
LOG-LOC-for.example    fall-IRR    PRESIMPF trip-throw-NONPAST-TOP  
‘*Jinarn-kijirni*. We call it *jinarn-kijirni* if you fall over on, say, a piece of  
wood, it trips one up.’

Were the highlighted *ka* in the final clause to really instantiate an *enclitic* we would have a good example of a type 1 *non-sister* clitic (unless a convincing case could be made for *ka* occupying its underlying syntactic position in (7), i.e. T<sup>0</sup>). Legate (2008) however, argues against an enclitic analysis of first position *ka* citing Laughren (2002), who characterizes the relevant examples as “involving a pause between the preceding material and *ka*” (Legate 2008:10). It thus appears that the sole exemplar of a regular type 1 non-sister clitic in the literature is under suspicion and should probably be excluded for non-sisters while admitted for sister clitics.

### 3.2.2 *Type 2: initial, before, proclitic*

This type also shows a strong division between sister and non-sister clitics and thus warrants special attention here. Non-sister clitics have been claimed to be unable to procliticize by Werle (2004, 2009). This claim will be tested by considering several candidates for non-sister type 2 clitics. Several of these will be rejected but at least one appears to be a true exemplar of this pattern. Werle’s generalization is thus widely

supported but not perfectly watertight, a fact which will be incorporated into the framework presented in chapter 4.

Exemplifying type 2 clitics on the level of the DP with sister clitics is trivial as it represents the unmarked case for prepositions and prepositional case markers. More interesting are examples on the clausal level. For this we can look at question markers, which can be thought of as taking clause level complements and thus still represent simple clitics when appearing on the edge of the clause. Four good examples are available in Austronesian. In two languages, Maranao (S. Philippines) and Seediq (Taiwan), the question marker can host 2P clitics (shown in italics), as seen in (8) and (9), respectively.<sup>5</sup> In the other two languages, Kulawi (Central Sulawesi) and Chamorro (Mariana Islands), the question markers appear in roughly the same position but cannot host 2P clitics, as shown in (10) and (11), respectively.

- (8) **Ba**=*ako=yka* di' ka-taw-i? *Maranao*  
 QM=1S.NOM=2S.GEN NEG NONV-know-DEP.LV  
 'Am I not known to you?' (McKaughan 1958:22)
- (9) **Yo**=*su* kula-un seedaq m<n>huma bulebun-ni? *Seediq*  
 QM=2S.GEN know-PV person <PRF>plant banana-this  
 'Do you know the person who planted these bananas?' (Aldridge 2002)
- (10) **Ba**=i-tudu tina=mu=*da=ko* *Kulawi*  
 QM=PV.RL-send mother=2S.GEN=EMPH=2S.NOM  
 'Are you sent by your mother?' (Adriani & Esser 1939:30)
- (11) **Kao**=patgon=*ña=hao* ädyu na ma'estra? *Chamorro*  
 QM=child-3S.GEN=2S.NOM that LNK teacher  
 'Are you the child of that teacher?' (Chung 2003)

---

<sup>5</sup> As a potential host for other clitics, these elements are only clitics themselves in that they cannot stand alone. Clitics which are able to host other clitics, although perhaps the marked case, are described in various places in the literature.



No clear examples of type 2 non-sister clitics have been brought to light in the literature. Among several candidates, Tondano (Minahasan; N. Sulawesi) is the least ambiguous attestation of such a clitic. As can be seen in Table 3.2, Tondano has quite a complex pronominal system, with three types of nominative pronoun. The first type is the free form, the only form which can be used on its own and also the form which is used for left-dislocated topics, as seen in (12)a. This form cannot be used in the post-predicate domain, as seen by (12)d. The second type is the reduced nominative which is used in the post-predicate domain as in (12)b, and can also be used as a prepredicate subject as in (12)c. The third type is the so-called topic concord form (Sneddon 1975), which is the form of interest as it appears to be a dependent clause-initial form, which can additionally double the subject (12)a-d.

**Table 3.2.** Tondano Pronominals

	FREE FORM	REDUCED NOM	TOPIC CONCORD	GENITIVE
1 <sup>st</sup> sing.	<i>niaku</i>	<i>aku</i>	<i>ku</i>	= <i>ku</i>
2 <sup>nd</sup> sing.	<i>niko ~ nikoo</i>	<i>ko ~ koo</i>	<i>ko</i>	= <i>mu</i>
3 <sup>rd</sup> sing.	<i>nisia</i>	<i>sia</i>	<i>si</i>	= <i>na</i>
1 <sup>st</sup> inc. pl.	<i>nikita</i>	<i>kita</i>	<i>kita ~ ta</i>	= <i>ta</i>
1 <sup>st</sup> excl. pl.	<i>nikey</i>	<i>key</i>	<i>key</i>	= <i>mey ~ =møy</i>
2 <sup>nd</sup> pl.	<i>nikow</i>	<i>kow</i>	<i>kow</i>	= <i>miow ~ =miu</i>
3 <sup>rd</sup> pl.	<i>nisea</i>	<i>sea</i>	<i>se</i>	= <i>nea</i>

- (12) a. **Niaku** **ku**=me-kaan                      b. **Ku**=me-kaan *Tondano*  
 1S      1S.NOM=AV-eat                      1S.NOM=AV-eat 1S.NOM  
 'I am eating.'
- c. **Aku**    **ku**=me-kaan                      d. \***Ku**=me-kaan    **niaku**  
 1S.NOM 1S.NOM=AV-eat                      1S.NOM=AV-eat 1s  
 'I am eating.'

Topic concord pronouns are obligatory with a post-predicate or non-overt subject, as demonstrated by (13). In the presence of a topicalized subject, however, they are



the material intervening between the subject clitic and the predicate head forms some kind of clitic complex (as commonly found in Oceanic languages) because the intervening negator *ra'i* is a free element and also hosts 2P clitics such as *mow* and *tu'u* in (16) (clitic host underlined). The Tondano case then appears to be a robust example of IP initial pronominal proclisis which shows that the grammar must allow type 2 non-sister clitics in principle.

### 3.2.3 *Type 3: initial, after, enclitic*

Type 3 is the canonical 2P clitic which has been amply exemplified in the preceding chapters. Note, however, that we have been heretofore only discussing pronominal and adverbial 2P clitics. In line with the goals of this chapter we must ask if this type is also attested with unambiguous morphosyntactic sisters. Besides adverbs and pronominals, definite markers have also often been used to exemplify type 3 clitics on the level of the noun phrase. Definiteness, however, is not a clear instance of a syntactic head but is rather more often understood to be a feature/interpretation associated with DP/NP (see Börjars 1998, Diesing 1992 and references therein). On the other hand, case markers and adpositions make unambiguous exemplars of morphosyntactic sisters, the critical difference being that these elements choose their complements selectively and are subcategorized for directly by higher heads, unlike the situation obtaining with definite NPs. Interestingly, there appear to be no good attestations of 2P clitic case markers or adpositions.

Marantz (1989:109), who observes the same gap in his discussion of Yagua type 1 case markers, notes that placement of such elements in NP-internal 2P would disrupt the government relation between the determiner and their governing verbs. Marantz's basic idea is that case markers need to be in some sense visible to higher

heads. This seems correct and will also serve here as a partial basis for the gap of type 3 case markers and adpositions.

Three potential counterexamples should be discussed here. The case of Ancient Greek *hyperbaton* which applied to boolean operators would be such a case under an analysis which treated those elements as syntactic heads. Similarly, conjunctions also take second position in West Greenlandic (Sadock 2003). Agbayani & Golston (to appear) argue convincingly however that this is not the correct analysis for these elements in Greek and according to the very simple criterion above of selectiveness and selectedness, boolean operators and other conjunctions clearly fail to qualify as heads.

The second potential counterexample are what Bowen (2004) describes as 2P complementizers in Bardi, an Australian language, specifically the morpheme *amb* which she glosses as COMP, as shown in (19).

- (19) Gorn=**amb** inin niyarra nganarlial *Bardi*  
good=COMP 3S.be taste 1S:PST:eat:REC.PST  
'Because it tastes good, I ate it.' (Bowern 2009, p.c.)

Upon further inspection of the data in Bowern (2004), it seems that the *amb* does not necessarily indicate subordination, but rather simple consecutive action, and thus appears to be most often translated into English as 'then'. If the morpheme in question only indicates sequential marking rather than the presence of a CP constituent, then it does not stand as a counterexample as sequential semantics can easily be considered as adverbial and is not selected by a higher head, as I am assuming true complementizers are. As it turns out, Bowern (2004:51-53) discusses several general problems with clausal subordination in Bardi, suggesting that true subordination may not exist at all in the language (cf. Hale 1976 for a similar claim regarding Warlpiri). The Bardi 2P clitics nonetheless deserve further inspection

especially considering Bower's claims that these derive from case markers historically (as cited in Bower 2004:51).

Finally, there is the more obvious question of V2 languages, as commonly found in Germanic, but also in Kashmiri and several other languages discussed in Anderson (2005, chap. 7). These are technically not of immediate concern as the elements in question are not claimed to be clitics. Nonetheless, V2 phenomena and 2P clitics have been given a similar analysis by Legendre and Anderson, following Wackernagel's observation that they may share a common origin. Although there may be similarities, there are some crucial differences which disqualify them from constituting a counterexample to our generalization. Namely, what is in 2P in V2 languages is not the head of VP, but rather the tensed element. The difference is that while elements like V are undoubtedly in a head-complement relationship with their objects, Tense need not be analyzed as a syntactic head along the same lines. Tense, or more specifically, the features [+T] or [+fin], do not take lexical complements. In fact, there is no agreement in the literature on what the complement of  $T^0$  should be or if it should be uniform across contexts and languages (cf. Pollock 1989, fn. 19).<sup>6</sup> Contra Pollock (1989) and much following work, I take this to signal that the evidence for T as a bona fide syntactic head is fundamentally weaker than that for V and thus subject to far more theory internal considerations.<sup>7</sup> Furthermore, V2 languages appear to display a marked difference with 2P clitics in languages like Tagalog and Chamorro in that the element in 2P never breaks up syntactic constituents. A standard syntactic raising account of a verbal element to T (and then possibly higher) is thus not at all at odds with the current proposal. Functional elements which have no selectional

---

<sup>6</sup> The relationship of  $T^0$  to subjecthood and nominative case has also been questioned (see McFadden 2004 for a thorough discussion), although this connection has a much firmer hold in current theorizing.

<sup>7</sup> It should thus of no surprise that Tense is treated in a very different and non head-like way in such theories as LFG (Bresnan 2001), HPSG (Sag, Wasow & Bender 2003) and Categorical Grammars (e.g. Steedman 2000).

relationship to lexical elements are thus interpreted here as potentially consisting merely of adjoined features, barring syntactic evidence to the contrary.<sup>8</sup> The treatment of putative tense/aspect elements as features rather than syntactic heads also holds for 2P “auxiliary” elements in Slavic and Australian languages which have similar tense/aspect/mood semantics.

It is thus concluded that type 3 clitics, the classic 2P type, while very commonly attested with pronominals and adverbials, are not found with case markers, adpositions, or other unambiguous syntactic heads.

#### 3.2.4 *Type 4: initial, after, proclitic*

Type 4 is now generally agreed upon to be unattested for any type of clitic and does not require much comment. Klavans sole evidence for this type was based on Steele 1977 which in turn relied on a grammar her interpretation of an early Spanish grammar of Tepecano, an extinct Uto-Aztecan language. The morpheme in question is *an*, glossed simply as CLITIC.PRON in (20).

- (20) ndedos      n=**an**=ahohoinda.  
 my.fingers    INTRODUCER=CLITIC.PRON=will.shake.them  
 ‘I will shake my fingers.’      (Steele 1977, Klavans 1985)

It is claimed to be positioned in 2P but to lean rightwards. The data is very incomplete however and various arguments against interpreting it as a type 4 clitic are presented by Marantz (1988), Sadock (1991), Halpern (1995) and Billings (2004).

In the absence of any further evidence that procliticization is possible in 2P, we concur with the above authors that it is categorically ruled out by the grammar.

---

<sup>8</sup> This is similar to Anderson’s (2005) treatment of putative inflectional heads as phrasal affixes.

### 3.2.5 Type 5: final, before, enclitic

Type 5 clitics present another controversial example which had been proposed by Klavans and consequently dismissed by later analysts on empirical grounds. Type 5 clitics are penultimate enclitics and were argued by Klavans to exist in the form of Nghanhcara dative clitics on the basis of Smith & Johnson's (2000) description. The evidence presented for this is shown in (21) with the clitic *ngu* DAT.3S. As can be seen in (21)b-f, regardless of the order of the clause internal constituents, the clitic always leans on what is to its left, while seemingly remaining in penultimate position.

- (21) a. *nhila pama-ng nhingu pukpe-wu ku?a wa:=ngu* *Nghanhcara*  
 he.NOM man-ERG him.DAT child-DAT dog give=DAT.3S  
 'The man gave the dog to the child.' (Smith & Johnson 2000)
- b. *nhila pama-ng nhingu pukpe-wu ku?a=ngu wa:.*  
 c. *nhila pama-ng ku?a nhingu pukpe-wu=ngu wa:.*  
 d. *nhila pama-ng ku?a pukpe-wu nhingu=ngu wa:.*  
 e. *ku?a nhingu pukpe-wu nhila pama-ng=ngu wa:.*  
 f. *ku?a nhingu pukpe-wu pama-ng nhila=ngu wa:.*

The problem, as pointed out by numerous commentators (see Billings 2004 and references therein), is that in all the examples the clitic is adjacent to the verb and thus probably instantiates verb-adjacent clisis with variable phonological attachment rather than penultimate clisis. The fact that Nghanhcara happens to be a verb final language is an independent fact, irrelevant to clitic position.

Another potential example is brought up by Nevis (1988:353), who claims that the Mansi conditional suffix *-ke* also instantiates Klavans' Type 5. But this morpheme is also required to be verb adjacent rather than simply picking out the penultimate position in the clause. Embick & Noyer (1999) discuss Mansi at length, arguing that it

is an in-situ complementizer (and hence does not qualify as a “ditropic” clitic in their terminology).

More recently, Peterson (2001) presents evidence in favor of a type 5 clitic from Ingush. Unfortunately, the data is not at all straightforward from a phonological perspective and further complicated by focus effects. Until this claim can be further substantiated, we take type 5 clitics to also be ruled out categorically.

### 3.2.6 *Type 6: final, before, proclitic*

The hypothetical type 6 clitic also occupies the penultimate position in its domain but instead of encliticizing, it procliticizes, an even more unlikely situation than that of type 5 clitics. Klavans’ (1985:113) admits to being “unable to find a clear example” but offers Sanskrit preverbs as a possibility. Sanskrit preverbs are, however, unsurprisingly, verb-adjacent and the fact that the verb is typically in final position in Sanskrit is an independent fact. The only positioning requirements of preverbs is thus that they are proclitic on the verb (Sadock 1991:72). With the exclusion of both 5 and 6, the mirror image of 2P can be eliminated. Clitics are thus never required to appear in penultimate position within their domain regardless of their phonological attachment.<sup>9</sup>

### 3.2.7 *Type 7: final, after, enclitic*

In arguing for the existence of type 7 clitics, Klavans again employs Spanish pronominals which are head adjacent and is thus unable to differentiate between the first and second parameters as the host is a unitary element. Subsequently, other

---

<sup>9</sup> Embick & Noyer (1999, 2001) discuss this positioning within the word domain for the Huave suffix *ay* which appears to prefer penultimate position when cooccurring with multiple suffixes. Not having access to additional data from Huave, I cannot comment on the facts here, although they are clearly relevant to the present discussion.



researchers have discounted the possibility of true domain final clitics (e.g. Billings 2004). It is not difficult, however, to exemplify type 7 for sister clitics and it is additionally possible to exemplify it for non-sister clitics. Taking question markers as syntactic heads on the clausal level, instances of type 7 clitics can be found in Mandarin, as in (22), and Ambon Malay, shown in (23), among numerous other languages of mainland East Asia.

(22) Qiaofen<sub>3s</sub> mai-le shenme **ne** *Mandarin*  
 Qiaofeng buy-ASP what QM  
 ‘What did Qiaofeng buy?’ (Cheng 1991:22)

(23) Aka<sub>3s</sub> ada ka=situ ka=situ=**ka**? *Ambon Malay*  
 3s EXT OBL=there OBL=there=QM  
 ‘Has it gone somewhere over there perhaps?’ (Van Minde 1997:261)

Non-sister pronominal clitics in this position are not common, but are attested clearly for Manggarai (Bima-Sumba, CMP; Flores, Indo.) and Irish (Chung & McCloskey 1987, Adger 1997, 2007, Elfner 2008).<sup>10</sup> As discussed by Arka & Kosmas (2005) and Arka (to appear), Manggarai subject clitics obligatorily attach to the right edge of the IP, where they can be separated from the verb by such things as prepositional phrases, as in (24), and oblique agents, as in (25). As can be observed from the examples below, these clitics regularly double full NP subjects.

(24) Hia pa’u eta mai bubu<sub>3s</sub> mbaru hitu=**i** *Manggarai*  
 3s fall above from top.roof house that=3S.NOM  
 ‘(S)he fell down from the top roof of that house.’ (Arka & Kosmas 2005:90)

(25) Latu<sub>3s</sub> hitu cero l=**aku**=**i** *Manggarai*  
 corn that fry by=1s=3S.NOM  
 ‘The corn is (being) fried by me’ (Arka & Kosmas 2005:95)

---

<sup>10</sup> Another possible example comes from Kichaga as discussed by Bresnan & Moshi (1990:166) which aligns strong pronouns to the right edge of VP.

In the case of Irish, the elements in question are generally referred to as weak object pronouns. Unlike the case of Manggarai, they cannot double full NP objects. The basic pattern is shown in (26) and (27). The unmarked position of arguments with full NPs is shown in the (a) examples while the unmarked order with weak pronominal objects is shown in the (b) examples. As can be seen from these latter examples, the object pronominal appears clause finally following all adjuncts.<sup>11</sup>

(26) a. Bhris sé an chathaoir leis an ord aréir *Irish*  
 broke he the chair with the hammer last.night  
 ‘He broke the chair with the hammer last night’ (Adger 2007:343)

b. Bhris sé leis an ord aréir í  
 broke he with the hammer last.night it.FEM  
 ‘He broke the chair with the hammer last night’

(27) a. Chunnaic mi an t-each anns a’ gharradh an dè *Irish*  
 saw I the horse in the garden yesterday  
 ‘I saw the horse in the garden yesterday.’ (Adger 2007:343)

b. Chunnaic mi anns a’ gharradh an dè e  
 saw I in the garden yesterday it.MSC  
 ‘I saw the horse in the garden yesterday.’

It is at this point unclear how many examples of pronominals following the Manggarai and Irish patterns can be found among the world’s languages but no others appear to have been taken into consideration in the clitic literature. Nonetheless, given the careful description by Arka & Kosmas and the rather well investigated case of

---

<sup>11</sup> The Irish facts appear considerably more complicated than the Manggarai ones due to rather wide variation and differential positioning based on syntactic context (small clauses, embedded clauses, etc.). Justice cannot be done to the facts discussed by the authors cited above. Suffice it to say, however, that an analysis which adjoins the relevant feature bundle either to the right edge of the clause or the left edge of the lowest phrase within the clause should be able to account for the facts as well as PF movement accounts such as that of Adger’s (1997, 2007), but demonstrating this must be left to further work.

Irish object pronominals, it can only be concluded that final non-sister clitics as well as sister clitics must be countenanced by the grammar.

### 3.2.8 *Type 8: final, after, proclitic*

On the clausal level, type 8 is the opposite of the putative Warlpiri clitic 1P enclitic. Both putative types are placed at the edge of a clause and are hosted by an element sitting outside that clause. Klavans tentatively exemplified this type with Classical Greek *ou* NEG but this analysis was shown by Sadock (1991) to be untenable. While the possibility of a final clitic leaning *outside the clause* to procliticize on following material can be justifiably discounted, it has not been sufficiently emphasized that the far more tenable situation of a *proclitic postposition* – that is, a head which takes its complement to the left but adjoins prosodically to the right – is also completely unattested. This is in contrast to the not uncommon case of enclitic prepositions. Any theory of clitic positioning and prosody must account for this asymmetry.

### 3.3 **Divergent patterns of sisters and non-sisters: a new typology**

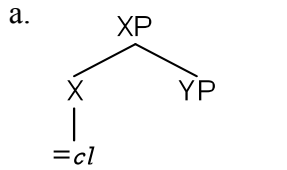
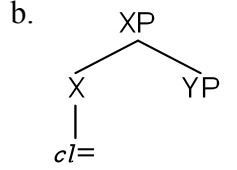
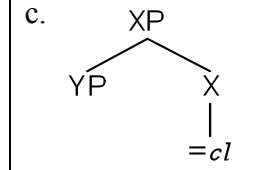
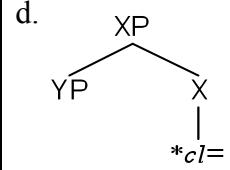
Klavans original typology has been reduced in much subsequent work as cited above. Some of these reductions have found support here while others have been counter-exemplified. In this section, we summarize the findings above in the context of the sister/non-sister dichotomy. We can begin by looking at those clitic types which have been ruled out altogether regardless of type, as shown in Table 3.3.. No convincing examples of Types 4, 5, 6 or 8 have surfaced in the literature, leaving only types 1, 2, 3 and 7. This distribution is a primary explanandum of any adequate theory of clitics.

**Table 3.3.** Summary of attested clitic types

Phrase	WP ]	[ XP				] [ YP		
Host position	A=1	2=B=3		4=...=5		6=C=7		8=D
	✓	✓	✓	x	x	x	✓	x

As we will see, the sister/non-sister distinction is key to understanding the distribution and prosodic parsing of clitics more generally. The typology of syntactic sister clitics is shown in Table 3.4. The node labeled X represents the head of XP with complement YP. Out of four logical possibilities, only three are found, with proclitic postpositional heads being the unattested type.

**Table 3.4.** Sister clitic typology

a.		b.		c.		d.	
E.g. Limos Kalinga case markers, K <sup>w</sup> ak <sup>w</sup> ala determiners	E.g. Tagalog case markers	E.g. Turkish postpositions (Bybee 2002)	—				

This absence has also been noted by Cysouw (2004), but he regards this as a result of probability rather than UG:

“In principle, examples of the mirror image phenomenon – postposed proclitics as in (1b) – are just as interesting, but I know of no convincing cases. I do not believe that there is any deep structural restriction at work here, but simply a strong cross-linguistic preference for clitics to be enclitic rather than proclitic, just as affixes show a strong preference for being suffixes rather than prefixes (cf. Halpern 1998:119). As proclitics are only rarely attested, and ditropic cliticisation is also a rare phenomenon, the combination of these two rare phenomena will be extremely rare.” (Cysouw 2004)

Problematic here is the claim that proclisis is ‘only rarely attested’. There is, on the contrary, a wealth of evidence that proclisis is very common cross-linguistically, especially if we (justifiably) include all examples of prepositions and preposed case

markers which do not satisfy prosodic wordhood requirements on their own. This false impression arises from the tendency of linguists to focus on enclisis due to its often more dramatic phonological interactions with the host and enclitic, a generalization which in turn is rooted in the stronger integration of left-leaning material. In sum, there is no good basis for claiming that the absence of proclitic postpositions to be a mere consequence of probabilities. Rather, it should be ruled out by universal principles.

Turning now to the typology of morphosyntactic non-sister clitics, we find that the options are far more limited. Only three possibilities are clearly attested: type 2 (first position proclitic), type 3 (second position enclitic) and type 7 (final position enclitic). Type 2 is only marginally attested with clear non-sisters, e.g. pronominal clitics at the left edge of TP, as in Tondano. Type 3, the classic 2P clitic, is amply attested for non-sisters, in Austronesian languages and beyond. Type 7 is found commonly for adverbial clitics in head initial languages, but is relatively uncommon for pronominal clitics which are unambiguously detached from the  $\nu$ P domain, as in Manggarai and Irish. The non-sister clitic types are summarized and compared to sister clitics in Table 3.5. Type 2 for non-sisters is marked with “~” indicating its marked nature and the fact that only one example has been brought to light.

**Table 3.5.** Attested types of sister versus non-sister clitics

Phrase	WP ]	[ XP ]						[ YP
Host position	A=1	2=B=3	4=...=5			6=C=7	8=D	
<i>sister</i>	✓	✓	✗	✗	✗	✗	✓	✗
<i>non-sister</i>	✗	~	✓	✗	✗	✗	✓	✗

Three interesting generalizations can be made from Table 3.. Proclisis is commonplace for sister clitics but extremely marginal for non-sister clitics (as also emphasized by Werle (2004, 2009) in different terms). This can be distilled as the following:

### Prosodic Asymmetry Generalization

Rightwards prosodic attachment is coerced by direct morphosyntactic constituency.

We also see from Table 3. that while sister clitics differ from non-sisters in freely allowing proclisis to their host, non-sister clitics stand out in being able to appear in 2P (position 3). For present purposes, this is best characterized descriptively as a constraint against displacement of syntactic heads to 2P:

### Syntactic Displacement Constraint

Unambiguous heads of phrases are never displaced to 2P

As noted above, there are also four complete gaps in the typology: types 4, 5, 6 and 8. Three of these, 4, 6 and 8, are proclitic on an element which is not a morphosyntactic sister and thus fall under the Prosodic Asymmetry Generalization. Type 5 clitics (penultimate enclitics), however, are parsed with material to their left and thus do not violate Prosodic Asymmetry Generalization. If type 5 represents misalignment from the right edge of XP, there must be an asymmetry regarding left and right edges such that only elements at the left edge are motivated to misalign. This results naturally from our earlier \*WEAKSTART constraint, which capitalized on the prominence of initial positions and was thus claimed to be inherently asymmetric. Pending the forthcoming analysis, the relevant descriptive generalization is captured by the following:

### Edge Asymmetry Generalization

Syntactic displacement of a clitic from the edge of its host only occurs on the host's left boundary

Focusing on the complementarity between clitic types 1 and 3, we can note that both misalignments from the selected edge (i.e. position 2) result in encliticization

(i.e. the unmarked direction of attachment). The two types of clitics achieve this through different means; sister clitics, by prosodic misalignment and non-sister clitics, (preferably) by syntactic misalignment. A satisfactory theory of clitic position and prosodification must account for all three of these generalizations.

In the following subsection we briefly compare one previous approach to this problem and evaluate its potential to capture the typology described here.

### 3.3.1 *Clitic typology and previous approaches*

An extremely wide range of analyses make reference to Inkelas' (1990) theory of Prosodic Subcategorization (henceforth PS) in some fashion. PS allows all morphemes to be specified in the lexicon as having prosodic requirements either to the left or to the right, as was suggested to be possible by Klavans' original study. The subcategorization frames are typically represented as in (28).

- (28)
- a. Enclitic to prosodic word:  $[[ \ ]_w \_ ]_w$
  - b. Enclitic to prosodic phrase:  $[[ \ ]_w \_ ]_w$
  - c. Proclitic to prosodic word:  $[ \_ [ \ ]_w ]_w$
  - d. Proclitic to prosodic phrase:  $[ \_ [ \ ]_w ]_w$

In theories which rely on PS for 2P phenomena, clitics which subcategorize for a prosodic word on their left must invert with following material to satisfy their requirement or must attach directly to the first prosodic unit within a larger containing unit (as in Chung 2003). On one hand, PS has greatly advanced our understanding of the role of prosody in clitic placement by allowing clitics to operate independently from syntax. On the other hand, because all cliticization is derived by PS frames in the lexicon, it is difficult, if not impossible, to make reference to syntactic structure and to motivate the asymmetries adduced above. In particular, as noted by Werle (2004, 2009) and Anderson (2005), the subcategorization frames for proclitic attachment in

(28)c and d are completely independent of syntax and thus mistakenly predict that morphemes can procliticize regardless of whether they are sister or non-sister clitics. As the present survey further offers strong evidence against Klavans' claim that phonology and syntax are completely independent factors in clitic positioning, our first order of business in building a more explanatorily adequate theory of clisis is the integration of this dependence in a natural manner.

Sadock (1991), Halpern (1995) and Anderson (1996) all observe the near unattested status of non-sister, domain initial proclitics but only Sadock attempts to incorporate this observation into his theory. Sadock proposes the following "interface constraint": "A lexeme that violates a homomorphism constraint is a suffix" (Sadock 1991:71). Sadock's homomorphism constraint militates against mismatches between syntactic constituency and linear order and thus the interface constraint essentially states that displaced clitics are enclitics.<sup>12</sup> Halpern (1995) acknowledges this observation and articulates it in more detail:

"...it is interesting to note that the taxonomy of positions for special clitics is relatively restricted. In fact, most special clitics seem either to appear in second position of some domain or to be verbal clitics. (Sentential clitics in head-final languages are often final, but this is essentially equivalent to the case of verbal clitics.) One thing which is notably missing is special clitics which are routinely initial (and proclitic)." (Halpern 1995:182-183)

However, Halpern seems to take a step backwards by not attempting to derive the absence at all in his theory. Rather, he attributes the unattested nature of "special

---

<sup>12</sup> Sadock is equivocal on whether or not the original observation is meaningful or not:

"Whether it is necessary to make this an absolute constraint on the system is unclear to me. If, for example, prefixes are simply rarer than suffixes, and mismatched lexemes are simply rarer than those whose structural positions in two autonomous representations are compatible, then these three types will simply be so uncommon that our limited data may not happen to contain any examples." (Sadock 1991:255 fn.8)



proclitics” to an alleged preference on the part of proclitics to avoid absolute initial position and on the part of enclitics to avoid absolute final position (Halpern 1995:183). Here, Halpern seems to have Tobbler-Mussafia effects in mind, where a pronominal clitic attaches to the left edge of a verb unless that verb is in some initial position, in which case it attaches to the right edge of the verb. This has however never been shown to be a regular characteristic of proclitics and has more importantly never been found to occur with syntactic heads (e.g. adpositions and case markers). Evidence that enclitics prefer to avoid absolute final position (however defined) has also been weak, generally referring to tendencies rather than rules (see Klavans 1994, Franks & King 2000 for examples).

Werle (2004) proposes that “prosodic markedness constraints force enclisis of most function words, but that interface constraints prefer that prosodic boundaries match syntactic boundaries, yielding proclisis of some words.” Taking Prosodic Clitic Theory (Selkirk 1995, 1996, Truckenbrodt 1995, 1999, Basri et al. 1998) as his base, Werle is able to derive the proclisis of prepositional heads by alignment constraints which enforce the edges of prosodic words to match up to lexical words. Just as in the present work, enclisis is understood by Werle as a default parsing with proclisis being coerced by syntactic structure. Unlike the present work, however, Werle (2004) depends on syntactic head movement to derive the correct input for the prosodic constraints. Prepositions must trigger movement of the heads of their complements and only once the two elements share a syntactic node can they be parsed together as a single prosodic word. A disadvantage of this is that it relies on otherwise unmotivated syntactic movement to derive a phonological effect.

Attempting more generally to reduce the arbitrary prosodic attachment permitted by Prosodic Subcategorization, Anderson (2005) emphasizes the importance of stray adjunction, arguing that the direction of stray adjunction within a given

language will determine how clitics are parsed prosodically. Anderson (2005:60) critically examines Nespor & Vogel's (1986) arguments for specifying the direction of prosodic affiliation for Greek clitics on the basis of the data in (29).

(29) a. O Dáskalos **mu**=to=ípe *Modern Greek*  
 ART teacher me=it=said  
 'The teacher said it to me.' (Anderson 2005:60 via Nespor & Vogel 1986)

b. O Dáskalos=**mu** to=ípe *Modern Greek*  
 ART teacher=me it=said  
 'My teacher said it.' (ibid.)

Because, the same clitic appears to attach both to the left and to the right depending on its function, Nespor & Vogel argue that the prosodic affiliation must be specified in the lexicon. But as Anderson shows, in this case, as in other cases, lexical specification is totally unnecessary as prosodic affiliation can be predicted by morphosyntactic affiliation. The first person DAT/GEN clitic in (29)a is an argument of the verb and thus morphologically affiliated with it while in (29)b it is a modifier on the noun and therefore affiliated with that constituent in the syntax. All that is necessary then is a principle which maps morphosyntactic constituents to prosodic ones and the facts are derived without stipulation. Nonetheless, by not distinguishing clearly between syntactic sisters and non-sisters, certain clitics within the same language must still be specified diacritically for which direction they are parsed in by stray adjunction. The problem is clear in his treatment of Bulgarian:

“In Bulgarian, for instance, a few sentential clitics occur initially (*ne*, *šti*) although most do not. It is not hard to show that Stray Adjunction works in both directions in this language, aligning syntactic and prosodic boundaries where possible. The difference between *ne*, *šti*, and other clitics can be ascribed to the fact that while all are subject to **LeftMost** (*cl<sub>i</sub>*), the **LeftMost** constraint requirement of *ne* and *šti*

dominates the alignment constraint (which we might treat as **LeftEdgeFaith** (CP) or the like), while those of the others do not.” (Anderson 2005:141)

On the present theory, the difference between *ne* NEG, *šte* FUT.AUX is that they are true syntactic heads while the other clitics are best considered non-sister clitics. Support from this comes in the fact that both *šte* and *ne* surface in the phrase structure where we expect them to and thus cannot be considered “special clitics”.<sup>13</sup> *Šte* furthermore alternates with non-clitic conjugated allomorphs in the future preterite and renarrated forms and can also serve as a phonological host for verb adjacent pronominal enclitics (Franks & King 2000:59). It is thus not that *šte* and *ne* are subject to higher ranking LEFTMOST constraints but rather that they are syntactic heads which take clause level complements and are thus subject to the Syntactic Displacement Constraint, as given above.<sup>14</sup> They need not be aligned by a LEFTMOST constraint at all, as assumed by Anderson, but can rather surface as the regular terminal nodes of syntactic projections. They are only exceptional in their phonological dependence but the direction of their dependence can be predicted by syntactic structure, as heads will cliticize to their complements in the unmarked case.

The last approach towards clitic typology to be reviewed here is that of Billings (2004), who derives clitic position by the following three constraints:

---

<sup>13</sup> For instance, *ne* NEG > *šte* FUT.AUX reflects the NEG>T order argued by Zanuttini (1993) to be universally available.

<sup>14</sup> Some arguments have been put forth by Rivero (1993), Rudin (1993), King (1996), for treating *li* as a C<sup>0</sup> element in Bulgarian and Macedonian. Rudin et al (1999) summarize these as the following: (i) *li* has an interrogative clause typing function and (ii) *li* does not cooccur with other complementizers. Assuming the correctness of Rudin et al’s thorough analysis of Bulgarian *li* as attaching to the first prosodic word in its domain, *li* appears to contraindicate the Syntactic Displacement Constraint as a syntactic head displaced to 2P. Interestingly, Rudin et al (1999) argue on independent grounds that *li* does not take 2P within a clausal domain but rather within the C<sup>0</sup> node itself after multiple head adjunction creates a verb complex in C<sup>0</sup>. Although further investigation is necessary, this characterization is fully compatible with the generalizations proposed here as *li* remains in-situ within C<sup>0</sup> and is not syntactically displaced.

SCOPE – Elements precede the domain over which they have scope

ALIGN (clause, L; intonation phrase, L) – A clause’s leading edge must coincide with the leading edge of an intonation phrase

SUFFIX – Morphemes marked as suffixes must follow some PWd

The three constraints above conspire to produce 2P clitics but this requires several assumptions. It is assumed that pronominals take scope over an entire sentence; scope must be realized overtly; and scope can only be realized by movement to or appearance at the left edge of the clause (Billings 2004:17). Thus, SCOPE, with its putative semantic motivation, does the job of the morphologically oriented LEFTMOST in Anderson’s theory. It is not made clear however in what sense clitics take scope over other clause internal material.<sup>15</sup>

The ALIGN constraint is unexceptional and is accepted by all analysts in one form or another. Together with the SUFFIX constraint it derives 2P phenomena in disallowing a prosodic word dependency across an intonation phrase. For a [+SUFFIX] clitic to satisfy the SUFFIX constraint it must be displaced away from the edge. However, the existence of SUFFIX can derive the Edge Asymmetry only if there is no mirror image constraint PREFIX, but as formulated above, there is no reason to believe why SUFFIX would be asymmetric as it is based on the prosodic needs of the clitics themselves rather than the prominent nature of initial positions. Additionally, just as with Prosodic Subcategorization, it must be assumed that 2P clitics are prosodically dependent despite contrary evidence.<sup>16</sup> Billings’ theory also has trouble with domain final clitics as the SCOPE constraint is claimed to be asymmetric. He must thus derive

---

<sup>15</sup> Uriagereka (1995) could provide one possible approach to this question. As referential elements, clitics are associated with point of view and must thus move to a corresponding functional projection dubbed FP by Uriagereka.

<sup>16</sup> Billings parries these arguments by claiming that Tagalog pronominal clitics are in fact head adjacent and not of the phrasal type, as generally claimed. But this position is not supported by the facts, as will be shown in chap 6, and thus presents a problem for a theory which requires all 2P elements to be prosodically dependent.

final clitics like English genitive 's via INTEGRITY constraints over the entire phrase to which they attach. Unfortunately, these INTEGRITY constraints are only relevant for the clitic in question and must be ignored by the rest of the grammar. Finally, Billings' theory does no better than the others reviewed above regarding the Prosodic Asymmetry Generalization and the Syntactic Displacement Constraint as it does not make reference to the crucial dichotomy between sister and non-sister clitics.

### 3.3.2 *Prolegomena to a new account of clitic typology*

The pivotal distinction between sisters and non-sisters is understood here as a direct reflex of two distinct syntactic operations which collaborate to build linguistic structure. Linguistic structure, as understood here, subsumes both syntactic structure proper as understood in traditional terms, i.e. the nodes of tree structure, as well as elements which I claim here possess no terminal node in the syntactic tree but are rather the result of spelling out adjoined feature bundles. Concretely, the two operations can be understood as two flavors of MERGE, one which operates over terminal nodes to build phrases and another which selects feature bundles and adjoins them to the edges of existing structure. Following the introduction presented in §2.4.1, these operations are defined as Merge Terminal and Merge [F(eature)], which for ease of exposition, will be referred to as *Concatenate* and *Adjoin*:<sup>17</sup>

*Concatenate* (Merge Terminal) –

An element is merged as a syntactic terminal with a complement to form a larger syntactic object

---

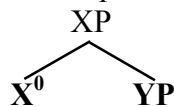
<sup>17</sup> Lex refers to both lexical material proper and functional items and is only meant to distinguish this operation from that which merges feature bundles. I take no stand here on whether elements are merged as fully formed words as in Chomsky 1995 or if words are built in the syntax as in Baker (1985), Halle & Marantz (1993) and much following work.

*Adjoin* (Merge [F]eature) –

A feature bundle is merged directly to an edge of XP/X<sup>0</sup> (without a terminal node)

The output of these two operations, shown schematically in (30), crucially differs when it comes to the parsing of syntactic structure into prosodic structure. When the phonology aligns a prosodic phrase to of the output of the ordinary Merge operation, *Concatenate*, both elements will naturally be included within the edges of the prosodic phrase, as represented trivially in (31)a. However, when the same operation applies to a structure formed by *Adjoin*, the feature bundle will be left outside the edge of the prosodic phrase corresponding to XP, as shown in (31)b. Without any “repair”, the feature bundle will be spelled out in a vacuum, a generally unacceptable outcome following the common assumption that all material must be somehow parsed prosodically (Selkirk 1995 *inter alia*).

(30) a. Output of *Concatenate* {X<sup>0</sup>, YP}  
(Syntactic complementation)



b. Output of *Adjoin* {F, XP}  
(Phrasal clisis)



(31) a. Synt: XP[X<sup>0</sup> YP]  
Phon: (X YP)

b. Synt: F XP[...]  
Phon: F (XP)

The most typologically common “repair” for the state of affairs in (31)b is to spell out the feature bundle in 2P rather than at its adjunction site. In this way, a crisp alignment between the left edge of the prosodic and syntactic phrase is maintained. Other options are possible, however, and given what I will argue to be a relatively unlikely ranking, the prosodic phrase can also be stretched or projected recursively to include the spell out of the feature bundle at its adjunction site. This prosodically motivated displacement does not apply to bona fide syntactic heads, which are

necessarily merged to their complements via *Concatenate*. This is because syntactic material can only be displaced by syntactic movement whereas adjoined features may be subject to displacement of a purely surface nature.<sup>18</sup> As will be shown in detail in the following chapter, the phonological pressure on bona fide syntactic heads is also less than it is on adjoined material because the natural prosodic parsing of syntactic phrases subsumes both syntactic heads and their complements. Nonetheless, as seen in the typology, syntactic heads can be phonologically disassociated from their complements when they are parsed as type 1 clitics, that is, as domain initial enclitics. Weak functional elements do not project prosodic word heads and are thus subject to the strictures of \*WEAKSTART as argued in chap. 2. Prepositional functional elements which are anchored to the tree structure can satisfy both their syntactic linearization requirements and \*WEAKSTART by staying in-situ and leaning to the left. For reasons of locality to be discussed in chap.5, this option is ruled out for adjoined features.

The current framework thus offers the following explanations for the three generalizations which have been distilled from the typological evidence:

Edge Asymmetry Generalization

Syntactic displacement of a clitic from the edge of its host only occurs on the host's left boundary

**Explanation:** Displacement is prosodically motivated and \*WEAKSTART is inherently asymmetric due to the unique prominence associated with domain initial positions.

Prosodic Asymmetry Generalization

Rightwards prosodic attachment is coerced by direct morphosyntactic constituency.

---

<sup>18</sup> Recall, however, the case of late merged adjuncts discussed in §2.4.1. This represents an exceptional case of syntactic structure being merged at PF “in between nodes”.

**Explanation:** Only bona fide syntactic material is visible to the constraints which align prosodic structure to syntactic structure. All else being equal, \*WEAKSTART will always prefer parsing orphan material with material to the right rather than material to the left.

#### Syntactic Displacement Constraint

Unambiguous heads of phrases are never displaced to 2P

**Explanation:** Displacement of bona fide syntactic material requires bona fide syntactic movement and in the case of downwards head movement to 2P the required movement is disallowed on general grounds (cf. Chomsky 1991 *inter alia*).

The last explanation requires some further explanation. While it remains possible that circumscribed types of lowering are admissible (see Richards 2004 and Chung 1998 for arguments to this effect), freely allowing lowering in order to capture 2P effects, as in Embick & Noyer 2001 (see §2.3.4), is both undesirable on theoretical and typological grounds as it overgenerates. The absence of lowering in syntax, which has been amply argued for in the literature on independent grounds, can thus also be seen to account nicely for the fact that unambiguous syntactic heads such as case markers and adpositions are never found in 2P. Rather, it is only feature bundles which are merged via *Adjoin* and thus not subject to the strictures of syntactic movement which may be displaced more freely by prosodic factors such as \*WEAKSTART.



### 3.4 Conclusion

In this chapter, we have reviewed the clitic typology first proposed by Klavans (1980) and, with previous researchers, eliminated some of the types that did not find support on further scrutiny. The primary contribution of the present chapter has been to argue that the options for clitic positioning and parsing are highly circumscribed based on the morphosyntactic nature of the clitic element itself. Several clitic types are only attested by syntactic heads while others types are only attested by elements with adverbial and phi-feature content. For bona fide syntactic heads, 2P is unavailable, and for adverbial and phi-feature elements, proclisis is highly marked and displacement is preferable. An explanation for this distribution was sketched in the syntax-phonology interface. It is proposed that the grammatical architecture allows two flavors of Merge; one which conforms with its traditional conception as an operation which combines syntactic terminals, and another which traffics solely in features, associating them with the edge of preexisting syntactic structure. When the syntactic structure is parsed, the constraints which wrap phonological phrases to syntactic phrases will treat adjoined feature bundles as orphans, putting them under far more pressure to be displaced at PF. Syntactic heads on the other hand must obey the rules of the syntax and thus behave differently when put under the same phonological pressures. Namely, these elements, when initial, may lean away from their complements in-situ rather than undergo displacement.

In the next chapter, the precise nature of the two Merge operations are examined in more detail and their consequences for prosodic parsing are formalized.

## CHAPTER 4: TWO TYPES OF STRUCTURE BUILDING

### 4.1 Introduction

In the previous chapter we saw that there is a significant difference in the way sister clitics and non-sister clitics are positioned and parsed in the output. To review, sister clitics are firmly anchored in their syntactic position but can lean away from their domain on the left edge. Non-sister clitics on the other hand may invert on the left edge rather than leaning outside their domain. The typology arrived at in chap. 3 is accounted for here using a combination of prosodic constraints and structure building operations which are sensitive to the difference between lexical items and purely functional items subject to late insertion. The two operations and their interface with prosodic constraints is the main topic of the present chapter. The goal will be to derive the different attested clitic behaviors from the position and, ultimately, the content of clitics themselves through two operations:

*Concatenate* (Merge Terminal) –

An element is merged as a syntactic terminal with a complement to form a larger syntactic object

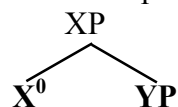
*Adjoin* (Merge Feature) –

A feature bundle is merged directly to an edge of XP/X<sup>0</sup> (without a terminal node)

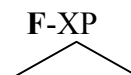
The difference in the output of these two operations is shown schematically in

(1) for head and feature initial structures.

(1) a. Output of *Concatenate* {X<sup>0</sup>, YP}  
(Syntactic complementation)



b. Output of *Adjoin* {F, XP}  
(Phrasal clisis)



In (1)a we see the output of *Concatenate*.  $X^0$  is merged with a complement phrase YP and the category label of the head is projected to the mother node as XP. In (1)b, a feature, F, is merged via *Adjoin* to the left edge of a phrase XP and in (1)c to the left edge of a head. Importantly, adjoined features are not properly included in the syntactic structure to which they are merged.

It cannot be the case that all elements are free to enter the syntax via either *Adjoin* or *Concatenate*. Clearly, bona fide lexical items, as opposed to functional items cannot be subject to *Adjoin* as it only applies to morphosyntactic features by definition. The converse, however, that all elements which consist solely of morphological features cannot enter the structure via *Concatenate*, does not hold. Functional heads, for instance, are typically not considered to contain lexical content. Nonetheless, they select for a phrasal complement and this selection requirement can only be saturated within a syntactic configuration. A functional head with a syntactic selection requirement is only able to satisfy that requirement within syntactic structure proper. Such items will thus necessarily be merged via *Concatenate*. On the other hand, pronominals, and arguably, many adverbials, are made up solely of morphological features. In principle, these elements have the choice of being merged as terminal nodes within the syntactic structure or as feature bundles adjoined to the structure. This choice has consequences for the form and behavior of the spelled out elements (cf. Everett 1996, as discussed in §2.4.1). What determines how a language merges a given set of features is beyond the scope of the present work. One possibility which can be entertained is that economy principles prefer *Adjoin* over *Concatenate*, as the former involves less structure. It would then only be the nature of a feature bundle in a given language which determines whether or not *Concatenate* is necessary, or alternatively, the features of selecting functional structure. If this is the case, we can

maintain the basic implication of Cardinaletti & Starke's (1994/1999) tripartite distinction for economy accounts of grammaticalization (cf. van Geldern 2004 *inter alia*) without all the unwelcome consequences of treating different pronominal types as different sized constituents). The structural reduction involved in grammaticalization would have its source in the locus and method of attachment rather than in X' categorial characteristics. More concretely, the grammaticalization cline witnessed with pronominal elements would correspond to the processes in (2).

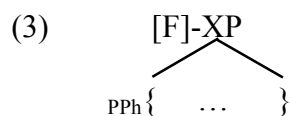
- (2) *Concatenate* {X, Y} > *Adjoin* {X, YP} > *Adjoin* {X, Y<sup>0</sup>}  
 Lexical pronouns            Phrasal/2P clitics            Head-adjacent clitics/affixes

The full implications of this suggestion cannot be worked out here. For present purposes, I assume economy of representation prefers *Adjoin* over *Concatenate* for purely functional elements but that selectional restrictions, among other factors, very often forces the use of *Concatenate*.

In the following we see how adjoined features are parsed prosodically and why they tend to appear in 2P.

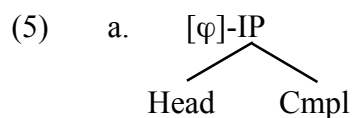
#### 4.2 Feature adjunction

When features are merged to a phrasal category via *Adjoin* they crucially lie outside the domain of that phrase. This has clear consequences for prosodic phrasing: given the alignment of prosodic phrases to syntactic phrases, XP-adjoined features will be spelled out in a prosodic vacuum, as illustrated in (3).



Assuming exhaustive parsing of all phonological material into the prosodic hierarchy (Selkirk 1995), these morphemes must either be phrased as adjuncts to an adjacent prosodic phrase or project their own prosodic phrase. Consider the parsing of the Tagalog disyllabic clitics below, which, as functional elements, may project prosodic words but not prosodic word heads (cf. the discussion of Zec 2005 in §2.4.3). As shown in (4), no parsing of the clitics in initial position is grammatical in Tagalog. Schematically, then, the morphosyntactic representation in (5)a for Tagalog, cannot be “rescued” by any of the prosodic mappings shown in (5)b.

- (4) \*niya táyo na-kítà-Ø  
 3S.GEN 1+2P.NOM NVL.BEG-see-PV  
 (For, ‘She saw us.’)



- b. Prosody (i): \*<sub>PPh</sub>[PWd PWd <sub>PPh</sub>[**PWd<sub>HD</sub>**]] (Recursion)  
 Prosody (ii): \*<sub>PPh</sub>[PWd PWd <sub>PPh</sub>[**PWd<sub>HD</sub>**]] (Independent projection)  
 Prosody (iii): \*<sub>PPh</sub>[PWd PWd **PWd<sub>HD</sub>**] (Inclusion)  
 Syntax: MWd MWd <sub>XP</sub>[MWd<sub>Lex</sub>]

The problem with domain initial positioning of adjoined clitics, both in Tagalog and cross-linguistically, is that it yields a prosodic parsing which will always be militated against by several very commonplace prosodic constraints:

ALIGN (PPh<sup>max</sup>,L; XP,L)

The left edge of maximal prosodic phrase is aligned to the left edge of an XP

ALIGN (XP,L;PPh,L)

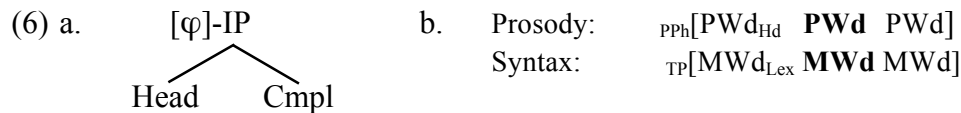
The left edge of an XP is aligned to the left edge of a prosodic phrase

\*WEAKSTART (PPh)

Violated by the configuration [**w (w)**..] in a prosodic phrase

\*STRUCTURE (PPh)  
Violated by a prosodic phrase in the output

Because weak pronominals, as purely functional items, cannot be prosodic word heads, they will violate \*WEAKSTART (PPh) and because adjunction to the right will involve misaligning the left edge of the maximal prosodic phrase from the left edge of a syntactic phrase, it will always violate ALIGN (PPh<sup>max</sup>,L; XP,L). The already beleaguered structure will furthermore either violate \*STRUCTURE (PPh) by the creation of an extra prosodic phrase layer, as in (i) and (ii) of (5)b, or ALIGN (XP,PPh) by wrapping the entire string within a single prosodic phrase, as in (iii) of (5)b. Projecting an independent prosodic phrase, as in (ii) of (5) above, only fares worse. The relatively “easy” solution to all of these difficulties is morphological misalignment, as in (6). In (6)a we see the person features aligned to the left edge of IP but misaligned from this edge in the output, shown in b.



In the unmarked case, feature bundles are spelled out at their adjunction site. However, just as word internal morphemes may be dislodged by phonological forces (Prince & Smolensky 1993), the spell out of feature bundles in syntactic structure can also be displaced by higher ranking prosodic constraints. This will incur a violation of the following general constraint:

ALIGN (Morph, [F])  
A morpheme in the output is aligned to the position of its corresponding feature bundle in the input

The crux of the present analysis lies in the fact that violation of ALIGN (Morph,[F]) satisfies several key constraints at once: (i) the clitics are not initial in PPh, satisfying \*WEAKSTART (PPh), (ii) they do not misalign a maximal prosodic phrase from an XP edge, satisfying ALIGN (PPh<sup>max</sup>,L; XP,L), and (iii) they allow the leading edge of the XP to be aligned to a prosodic phrase edge without prosodic phrase recursion, i.e., without needing to violate \*STRUCTURE (PPh). The violation profile of the three parsing alternatives in (5)b are shown in Tableau 4.1 below as candidates (a)-(c). The misaligned candidate which spells out the adjoined features after the lexical head thus violating ALIGN (Morph;[F]) once, is shown as (d). Note that one violation of ALIGN (Morph;[F]) satisfies multiple constraints. As we will see in the next section, the parsings in (a) and (b) are ruled out by an inviolable parsing principle and thus the only real competition is between (c) and (d); suggesting a difference of three constraint violations to one between the aligned and misaligned versions.

<b>Input:</b> [F] <sub>-IP</sub> [MWd <sub>Lex</sub> ]	*WS (PPh)	*STRUCT (PPh)	ALIGN-L (PPh <sup>max</sup> , XP)	ALIGN-L (XP;PPh)	ALIGN (Morph;[F])
a. PPh[ <i>ω ω</i> PPh[ <b>ω</b> <sub>Hd</sub> ]]	**	**	*		
b. PPh[ <i>ω ω</i> ] PPh[ <b>ω</b> <sub>Hd</sub> ]	**	**	*		
c. PPh[ <i>ω ω ω</i> <sub>Hd</sub> ]	**	*	*	*	
d. PPh[ <b>ω</b> <sub>Hd</sub> <i>ω ω</i> ]		*			*

**Tableau 4.1.** Violation profile for domain initial adjoined morphemes

So while it is not entirely impossible for the grammar to generate sentence-initial clitics – indeed, it appears necessary to account for cases such as Tondano from the previous chapter – it is challenging, as domination by any one of three *basic* constraints will result in misalignment. This accords well with the typological rarity of Klavans’ type 2 clisis to non-sisters.

For the sake of concreteness, we can compare the evaluation of the Tagalog facts above with the case of Tondano, which was argued to instantiate rare type 2

pronominal clisis to a clausal projection. For Tagalog we can simply fill in the relevant morphemes for the previous tableau. Any ranking in which \*WEAKSTART or ALIGN-L (PPh<sup>max</sup>; XP) are ranked sufficiently high as in Tableau 4.2 will derive the correct facts in the basic case.

<b>Input:</b> [ $\varphi_{Gen}$ ]-[ $\varphi_{Nom}$ ]-IP[ <i>nakítà</i> ] [3S]-[1+2P]-PV.NVL.BEG-see	*WS (PPh)	*STRUCT (PPh)	ALIGN-L (PPh <sup>max</sup> ; XP)	ALIGN-L (XP;PPh)	ALIGN (Morph; [F])
a. PPh[ <i>niya táyo</i> PPh[ <i>nakítà</i> ]]	*!*	**	*		
b. PPh[ <i>niya táyo</i> ] PPh[ <i>nakítà</i> ]	*!*	**	*		
c. PPh[ <i>niya táyo</i> <i>nakítà</i> ]	*!*	*	*	*	
d. $\wp$ PPh[ <i>nakítà niya táyo</i> ]		*			*

**Tableau 4.2.** Tagalog basic case

Recall the Tondano examples in (7) and (8) repeated here from chap. 3 showing that nominative features are aligned to and spelled out at a higher clausal projection (referred to here generically as IP). Tondano thus differs from Tagalog in disallowing misalignment of adjoined features.

(7) **Si**=ta'arəkan rai' wewe=ŋku *Tondano*  
 3S.NOM=almost NEG hit=1s.GEN  
 'I almost didn't hit him.'  
(Sneddon 1975:143)

(8) **Ku**=me-kaan **aku** *Tondano*  
 1S.NOM=AV-eat 1S.NOM  
 'I am eating.'

This simply means that ALIGN (Morph;[ $\varphi$ ]), the constraint which aligns the spell out of phi-features to their underlying position outranks all relevant prosodic constraints, as shown in Tableau 4.3. The choice between the three possible parsings of the edge-aligned clitic is decided by the relative ranking of these latter constraints. Because there is a disyllabic minimum on prosodic words in Tondano, parsing the monosyllabic *ku* into a prosodic phrase will either violate LAYEREDNESS (an inviolable



principle) or the prosodic word minimality constraint  $PWd \geq 2\sigma$ . Candidate (b) is thus rendered sub-optimal.<sup>1</sup>

<b>Input:</b> [1S.NOM] <sub>-IP</sub> [mekaan aku] AV.eat 1S.NOM	ALIGN (Morph; [φ])	ALIGN-L (XP;PPh)	*WS (PPh)	*STRUCT (PPh)	ALIGN-L (PPh <sup>max</sup> ; XP)	PWd ≥2σ
a. $\leftarrow$ PPh[ku PPh[mekaan aku]]			*	**	*	
b. PPh[ku] PPh[mekaan aku]			*	**	*	*!
c. PPh[ku mekaan aku]		*!	*	*	*	
d. PPh[mekaan ku aku]	*!			*		

**Tableau 4.3.** Tondano basic case

In the preceding cases the clitics in question were all in absolute domain initial position. The situation becomes more complicated in embedded contexts. In the next subsection, we examine embedded environments to further refine the analysis.

#### 4.2.1 *The status of complementizers and the Phase Correspondence Principle*

One of the more general challenges in understanding 2P clisis is locating the leftwards boundary of the clitic placement domain and deriving its position. In many cases, there appears to be no single leftward boundary which applies equally to all clitics. Different types of clitics may take different domains within the same language. Based on what we have seen so far, if morphological features are adjoined to the left edge of a phrase, any preceding material within the same tree should in principle be able to host the clitic in its adjunction site. Given a complementizer in  $C^0$ , for instance, a prosodically dependent morpheme should be able to cliticize to the left despite having its adjunction site to its right, as in (9). Even if the clitic in question is not prosodically dependent, the complementizer should be able to shield it from being initial in their prosodic phrase.

<sup>1</sup> In fact, we do not have sufficient evidence from Tondano to argue for parsing (a) over (c). This should just be taken as a hypothetical analysis in need of further support from phonological phenomena.

(9) CP[C=[F]-IP[...]]

This is, however, rarely the case, and thus something else must be at work here. An obvious problem specific to the case of complementizers could be in their own prosodic deficiency, as functional heads. Interestingly, there are several claims in the literature that complementizers and other conjunctions in certain languages can optionally host clitics but only when stressed. Browne (1974), Inkelas & Zec (1988), Zec & Inkelas (1990) note that 2P enclitics appear to induce stress on certain conjunctions when they play the role of host. In the current framework, these elements can be treated as optional prosodic word heads.<sup>2</sup> When they surface as heads they receive higher prominence and are viable clitic hosts, when they do not, they are unaccented and cannot host clitics. Thus, for features to be adjoined to CP does not guarantee their surfacing after the complementizer; the complementizer must also satisfy prosodic requirements for hosting clitics. But this is not the end of the story.

Morphosyntactic factors also appear to play a role and these can be teased apart in Tagalog, which treats pronominal clitics differently from adverbial clitics. We can compare the Tagalog adverbial 2P clitic *naman* SWITCH TOPIC with the pronominal *ka* 2S.NOM in the example in (10). The adverbial can attach to the complementizer, as seen in (10)a, while the pronominal appears after the first host in IP, which in this case is negation. For the pronominal to appear after the complementizer along with the adverbial is ungrammatical as shown in (10)b.

(10) a. Pára=**naman** hindi=**ka** mag-mukha=η gáya~gáya  
COMP=SWTCH NEG=2S.NOM AV-face=LNK ITER~imitate  
'So that you don't look like you're just imitating.'<sup>3</sup>

---

<sup>2</sup> See also Richardson (1997) and Anderson (2006:147) for discussion of optional clisis to complementizers in Czech and Taylor (1996:498) for Ancient Greek.

<sup>3</sup>From: [www.starmometer.com/2007/12/29/gma-shows-dominate-christmas-tv-ratings/?cp=5](http://www.starmometer.com/2007/12/29/gma-shows-dominate-christmas-tv-ratings/?cp=5)

b. \*Pára=ka=naman      hindi mag-mukha=ŋ gáya~gáya  
 COMP=2S.NOM=SWTCH NEG AV-face=LNK ITER~imitate

This demonstrates that the complementizer is a legitimate host from the point of view of the prosody but is simply out of range for pronominal arguments. The fact that monosyllabic conjunctions/complementizers in Tagalog such as *kuŋ* COMP, *na* COMP and *at* CONJ are unable to host either pronominals or adverbials underscores the need to tease apart issues of syntactic range from prosodic requirements. Our task then is to define “out of range” in terms of the syntax-prosody interface, rather than narrow prosody.

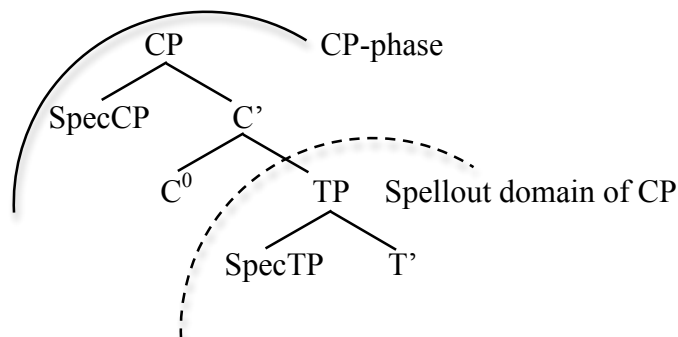
There is *prima facie* evidence that the relevant notion relates to the derivational phases of Chomsky (2000 *et seq*). Much recent work has uncovered possible prosodic consequences of phase theory (Kahnemuyipour 2003, Wagner 2005, Kratzer & Selkirk 2007) and thus effects on 2P clisis should not be surprising. In Chomsky (2000, 2001),  $\nu$ P and CP are identified as strong phases which are understood to correspond to semantic units (events and propositions, respectively) and “Spell-Out units”. Once a phase is complete and sent to Spell-Out, only its edge is visible to operations in the higher syntactic structure. What is relevant for our purposes is that the Spell-Out domain of CP is understood to be the complement of  $C^0$ ; merging of  $C^0$  triggers Spell-Out of TP. In the present theory, 2P pronominal clitics are the Spell-Out of features adjoined to TP. It would make sense then that these clitics must be instantiated within the same Spell-Out domain as their host and thus barred from cliticizing to C itself.<sup>4</sup> Adverbial clitics which have the CP layer as their semantic

---

<sup>4</sup> Note that this is also expected from a parsing perspective since the hearer will require extra-syntactic cues for properly parsing clitics which are not associated with a terminal node. There thus must be a stronger requirement for the prosodic parsing of adjoined clitics than of concatenated functional elements which possess a terminal node and project a phrase.

domain, e.g. Tagalog *naman* SWITCH TOPIC, adjoin to CP and can thus be spelled out in its domain.

(11)



I tentatively take the principle in (12) to determine the leftmost boundary of clitics crosslinguistically:<sup>5</sup>

(12)

Phase Correspondence Principle (PCP):

Features are spelled out within the phase which they are attached

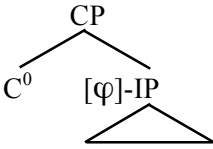
The PCP militates against certain clitic positions by ruling out certain prosodic parsings. In a basic embedded structure such as (13)a, with a post-complementizer clitic, only one of the parsings given in (13)b will be legitimate. In parsing (i), the clitic is clearly parsed outside of the minimal prosodic constituent as the phase thus violating the PCP. In (ii), all material is parsed together and thus technically satisfies

---

<sup>5</sup> The PCP, as stated in (12), enforces in representational terms a phonological effect of what is claimed to be a purely derivational aspect of the language faculty. If syntactic structure is sent to the phonological component cyclically (implemented here in OT) in phase sized chunks, then the PCP may be redundant; complementizers would simply be invisible at the time clitics are parsed. Unfortunately, the consequences of a cyclic phase based evaluation in an otherwise parallel OT grammar cannot be addressed here. See Elfner (2008) and Kratzer & Selkirk (2007) for some ideas on this.

The most immediate problem to be worked out in regard to this proposal is the fact that oblique wh-interrogatives are within the 2P domain of argument clitics in Tagalog and all other languages instantiating 2P argument clitics with which I am familiar. In Tagalog, the position of interrogatives is clearly below the complementizer, which is perhaps seated in ForceP (Rizzi 1997, see chapter 6 below). If the CP phase includes the position to which oblique interrogatives move than this position obviously should not be susceptible to any phase based locality generalizations. But it is unlikely that all languages which show 2P clisis to interrogatives differ significantly from more familiar languages in regard to cyclicity related conditions. This is a topic for further research.

the PCP but the edge of the Spell-Out domain is also generally understood to be aligned to a prosodic phrase boundary (see Kratzer & Selkirk 2007 and references therein). This flat parse includes multiple Spell-Out domains within a single prosodic phrase and thus violates the putative alignment between Spell-Out domains and prosodic phrase boundaries. Finally, (iii) satisfies both the PCP and correctly projects two corresponding prosodic phrases for the two Spell-Out domains. Assuming prosodic dependency on the part of the adjoined clitic, this final parse requires proclisis to following material. The conclusion then is that if a language adjoins pronominals to an IP layer, we should only expect post-complementizer positioning if the language also allows “1P” clisis, as in Tondano, as both cases require proclisis to a clausal non-sister projection.

- (13) a. 
- b. Prosody (i): \*<sub>PPh</sub>[PWd=**PWd** <sub>PPh</sub>[PWd<sub>Hd</sub>]]  
 Prosody (ii): ?<sub>PPh</sub>[PWd=**PWd** PWd<sub>Hd</sub>]  
 Prosody (iii): <sub>PPh</sub>[PWd <sub>PPh</sub>[**PWd**=PWd<sub>Hd</sub>]]  
 Syntax: <sub>CP</sub>[<sub>C<sup>0</sup></sub> **MWd** <sub>cl</sub> <sub>IP</sub>[<sub>V<sup>0</sup></sub> **MWd**<sub>Lex</sub>]]

The PCP may also provide an explanation for 2P clitics in other languages which cannot be hosted by CP material. Several cases from Germanic are discussed by Fontana (1996), as exemplified with the three varieties in (14)-(16). In each case, 2P pronominal clitics are unable to cliticize to the leftmost material in root contexts due to its syntactic position in CP. <sup>6</sup>

- (14) nu moete=**ne** onse vrouwe bewaren *Middle Dutch*  
 now must=**him** our lady save  
 ‘our lady must save him now’ (Van der Horst 1981, Fontana 1996)

<sup>6</sup> As discussed by the sources cited, the avoidance of CP does not appear to hold in subordinate clauses. That is, matrix clauses in Germanic pattern like Tagalog and subordinate clauses pattern like Tausug, discussed directly below. I do not speculate here on the root of the matrix/subordinate distinction here as several possibilities are available.

- (15) geschter het=**er**=nech=s zeigt Bernese (Swiss) German  
 yesterday has=he=to.you=it shown  
 ‘Yesterday he showed it to you.’ (Penner 1991, Fontana 1996)
- (16) Wat hat=**m** die Mutter gegeben German  
 what has=him the mother given  
 ‘What has mother given him?’ (Haverkort 1994, Fontana 1996)

There is however at least one language closely related to Tagalog which does allow 2P pronominal clitics to enciticize to complementizers. Billings & Kaufman (2004) compare the placement of pronominal clitics in Tagalog with Tausug, another Philippine language of the Bisayan subgroup (Zorc 1975/1977). Tausug generally allows clisis to complementizers and does not have a disyllabic requirement on clitic hosts (and perhaps prosodic words, more generally). Thus, the monosyllabic complementizer *baŋ* is seen to regularly host clitics, as in (17). The same facts hold for Jama Mapun (Sama, S. Phil.), as shown in (18).<sup>7,8</sup>

<sup>7</sup> Another difference obtains between matrix and embedded clauses due to the constraint ALIGN (PPh<sup>max</sup>, L; XP,L) which requires that every maximal prosodic phrase is aligned to a syntactic phrase. Compare the domain initial clitics in the IP structure in (i) with those in (ii), an embedded environment, where the element in bold designates the clitic. Assuming that the clitic realizes features adjoined to CP, it appears in its “underlying position” in both (i) and (ii). For it to surface as indicated in (i), ADJOIN would have to dominate both \*WEAKSTART (PPh), and ALIGN (PPh<sup>max</sup>, L; XP,L) since the maximal prosodic phrase is misaligned with the left edge of IP. The second constraint is not violated however in (ii), because the clitics are cushioned by a phrasal head, the complementizer; the maximal prosodic phrase can thus be aligned with the CP. This would suggest that, given CP-adjoined features, post-complementizer placement has fewer hurdles to overcome, even if the complementizer is not a prosodic word head.

(i) Prosody:  ${}_{PPh}[P\mathbf{W}d \quad P\mathbf{W}d_{Hd} \dots]$  (ii) Prosody:  ${}_{PPh}[P\mathbf{W}d \quad \mathbf{P}\mathbf{W}d \quad {}_{PPh}[P\mathbf{W}d_{Hd} \dots]]$   
 Syntax:  ${}_{IP}[M\mathbf{W}d_{Lex} \quad \mathbf{M}\mathbf{W}d \quad V^0]$  Syntax:  ${}_{CP}[M\mathbf{W}d \quad \mathbf{M}\mathbf{W}d \quad {}_{IP}[M\mathbf{W}d_{Lex} \quad V^0]]$

<sup>8</sup> Incidentally, the positioning pattern found with Tausug *baŋ* may have been borrowed along with the complementizer from one of the Sama languages which Tausug has been in close contact with (Pallesen 1985). The borrowing of functional elements together with parts of their associated syntax is also discussed by Kroch (1994) for the case of Persian complementizers borrowed into Hindi. Examples of post-verbal positioning can also be found in the presence of preceding complementizers in Tausug, as in (i), although this appears to be a marked option.

(i) Baŋ kabayaan=**mu** mag-bunu’ suŋ=na=kaw Tausug  
 if desire=2S.GEN AV-fight come=CMP=2S.NOM  
 ‘If you want to fight, come here.’ (Tawasil 1978)

- (17) baŋ=**kaw** k<um>aun magbalik... *Tausug*  
 if=2S.NOM <AV>eat again  
 ‘If you eat again...’ (Tawasil 1978:199)
- (18) boŋ=**ko** ya’ patagoŋ n-aŋis... *Jama Mapun*  
 if=2S.NOM NEG stop AV-cry  
 ‘If you don’t stop crying...’ (Collins, Collins & Hashim 2001)

The Tausug facts can be treated in two different ways. Either the PCP has been relaxed to allow pronominals to encliticize outside their phase, or argument features are simply adjoined higher in Tausug than they are in Tagalog, and other languages. The latter position is taken by Billings & Kaufman (2004), although both possibilities are in principle open.

In the next section we will look at the other side of the clitic typology: the behavior of morphosyntactic sister clitics merged via *Concatenate*. As we will see, the security of being a syntactic terminal and of being properly included within a phrase allows for more prosodic freedom, licensing the ability to be parsed prosodically with external material.

### 4.3 Concatenated clitics

Concatenated clitics are typically heads of their syntactic phrases, with the most unambiguous cases examined here consisting of adpositions, case markers and subordinators. In the unmarked case these clitics are parsed with their syntactic complements but when they are left headed (with a following complement), they may also be parsed with preceding material.

As observed in chap. 3, morphosyntactic sister clitics never undergo movement to 2P. The most straightforward interpretation of this fact, I claim, is that bona fide syntactic material can only undergo bona fide syntactic movement and that lowering to 2P does not constitute syntactic movement but rather morphological displacement at

PF.<sup>9</sup> This is supported by independent prohibitions on lowering (cf. Chomsky 1991) as well as ample evidence from Tagalog and Chamorro, reviewed in chap. 2, that 2P is not a position which can be derived by legitimate syntactic movement. The gap in the typology of 2P elements is thus derived from the architecture of the grammar, accounting for its exceptionless nature.

In this section we examine two phenomena which are characteristic of concatenated clitics: prosodic parsing with external material and proclisis.

#### 4.3.1 *Deriving enclitic prepositions*

We examine here how Klavans' type 1 clitics are handled by the present proposal. Data from Limos Kalinga, repeated below in (19) and (20), was employed in the previous chapter to exemplify these clitics. Kalinga has two allomorphs for the oblique marker, /si/ and /t/.<sup>10</sup> They are not completely predictable and therefore listed, but their choice is phonologically conditioned on the basis of the preceding segment. If the preceding segment is a vowel the case marker will encliticize as /t/ outside of its syntactic phrase, as in (19). Although Ferreirinho does not discuss this, we can further assume that the post-consonantal allomorph, /si/, shown in (20), is also enclitic, but we do not have segmental evidence for the direction of attachment (and nothing hinges on this decision).

- (19) Naŋ-anup dadit tagu=t bolok *Limos Kalinga*  
 AV.PRF-hunt PL.NOM person=OBL pig  
 'The people hunted pig.' (Ferreirinho 1993:12)

<sup>9</sup> Note that prosodically motivated syntactic movement is not ruled out by the current approach. Rather, it must be the case that such movement has at least some of the hallmarks of syntactic movement. See Zubizarreta 1998 for extensive discussion of prosodically motivated movement in Romance.

<sup>10</sup> Ferreirinho in fact claims that the alternation is between *si* and *ut* which is commonly reduced to *t* but all of her examples display this reduction. If this is correct, it would be surprising as a vowel-initial allomorph would be chosen precisely after a vowel-final word, creating hiatus.



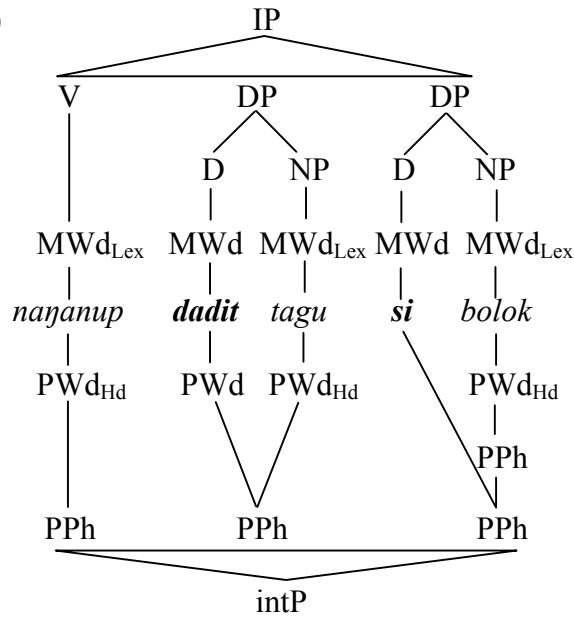
- (20) Mam-mula=ak **si** balat.  
AV-plant=1S.NOM OBL banana  
'I'm planting bananas.'

*Limos Kalinga*

(Ferreirinho 1993:82)

STRUCTURE, the constraint which is violated by all overt material, naturally prefers the reduced allomorph /t/. However, only heterosyllabic consonant clusters are allowed in Kalinga and thus /t/ can only surface adjacent to a vowel. We can assume that Kalinga is similar to Tagalog and many other Philippine languages in lacking truly vowel-initial roots. (Apparent vowel initial roots are in fact glottal stop initial.) This has the consequence that /t/ can only appear as a coda after a preceding open syllable if the parsing were to allow it. Below, we compare the option of obeying syntactic constituency with that of encliticizing to preceding material in the kind of dual representations familiar from Autolexical Theory (Sadock 1991). Syntactic constituency is shown in the branching structure above the sentence and prosodic constituency, below the sentence. A portion of the violations incurred by the structure as a whole, including correspondences between the syntax and prosody, are shown in the adjacent violation profile. Two primary candidates are shown in (21) and (22). In (21), the case markers *dadit* and *si* attach to their noun phrase complements and thus satisfy ALIGN (XP,PPh) which requires that every syntactic phrase be aligned to a prosodic phrase. This attachment, however, results in starting a syntactic phrase with a weak element, thereby violating \*WEAKSTART (PPh).

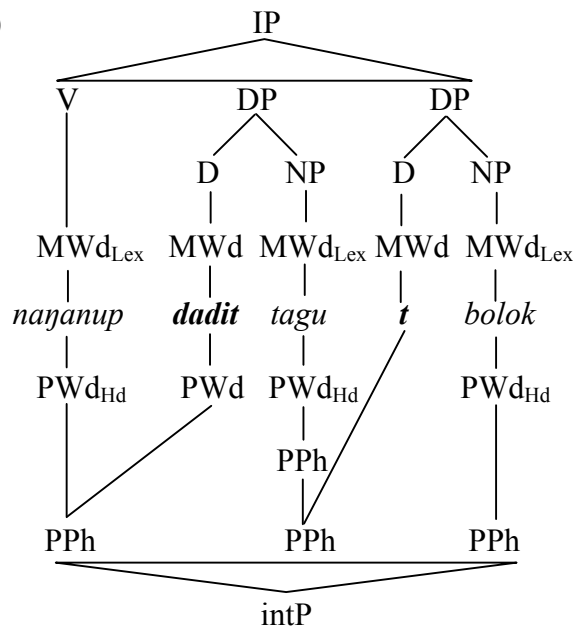
(21)



violation profile	
*WEAKSTART (PPh)	xx
ALIGN (XP,PPh)	✓

In (22), the attested parse with the allomorph /t/, the case markers break with their syntactic constituency, violating ALIGN (XP,PPh) in order to satisfy \*WEAKSTART (PPh).

(22)



violation profile	
*WEAKSTART (PPh)	✓
ALIGN (XP,PPh)	xx

The above violation profiles should suffice to show that the two types of prepositional markers can be derived by the following rankings (with the latter being correct for Kalinga):

ALIGN (XP,PPh) >> *WEAKSTART (PPh)	<i>prepositional proclitic</i>
*WEAKSTART (PPh) >> ALIGN (XP,PPh)	<i>prepositional enclitic</i>

Because \*WEAKSTART is inherently asymmetric, there is no combination of constraints in this theory which can force postpositional proclitics, one of the unattested variants in the typology. Note also that syntactic movement of the clitic, i.e. via head raising, could not possibly improve the situation. It would incur violations of STAY (Grimshaw 1997), the constraint which penalizes syntactic movement, only to place the head in a similar, phrase initial position. Head raising is thus never seen to repair violations of prosodic constraints.

#### 4.4 Head clitics and Feature clitics in K<sup>w</sup>ak<sup>w'</sup>ala

Before continuing to the interesting case of K<sup>w</sup>ak<sup>w'</sup>ala, it is worth reiterating the definitions of two key concepts:

**Inflectional morphology:** The spell out of functional features.

**Syntactic head:** An element with the ability to select a complement and to pass a categorial label to its mother node.

By the last definition, case markers must be considered heads since they may select a nominal complement and additionally must pass on their label to their mother node. This is not a theory-internal choice, but rather a logical necessity; the operations of phrasal syntax must be able to make reference to genitive phrases, nominative phrases and oblique phrases. Equally important, the phrasal syntax ignores the internal

make-up of these constituents, further emphasizing the head status of the case markers themselves. Spelled out functional features, on the other hand, never donate their label to their containing phrase. Rather, they contain inflectional information regarding a larger unit of reference or predication. The features which we are focusing on here are pronominal features associated with arguments. In accordance with our criteria for differentiating syntactic heads and features, they are selected and do not select. In many languages, they even appear obligatorily in the presence of a corresponding full NP argument.

The above division of elements into heads and features suggests an elegant solution to a particularly thorny problem in  $K^{w}ak^{w'}ala$  discussed by Anderson (2005).  $K^{w}ak^{w'}ala$  (Wakashan; American Northwest Coast) possesses two types of nominal clitics which appear simultaneously in DPs. The first element, mentioned earlier in the discussion of type (a) clitics, appears on the left edge of the determiner phrase but encliticizes to preceding material. The second class, referred to as post-nominal determiners by Anderson, appear in second position. This is exemplified in (23), where  $x'ux'da$  is the prepositional enclitic determiner and  $ix'$  is the 2P clitic.

- (23)  $dux'wida-s=x'-ux'da$   $guk^w=ix'$   $K^{w}ak^{w'}ala$   
 see-you-OBJ-DEM(2) house=DEM(2.VIS)  
 'Do you see this house (near 2<sup>nd</sup> person, visible)?' (Anderson 2005:102)

This is potentially problematic for a theory of clisis which treats all clitics as equals and requires 2P clitics to be misaligned from the left edge for prosodic reasons. If the first determiner can find a host to the left then the second determiner should as well. Anderson sums up the problem in relation to a PI-type analysis:

“...the Prosodic Inversion analysis cannot accommodate the  $K^{w}ak^{w'}ala$  second determiners. Assuming that these clitics are initially generated

in the left periphery of the DP, they would already be preceded by material that could support left attachment, without requiring any inversion. We can tell that the material preceding the left edge of the nominal is an appropriate host for clitics, because in fact the phrase-initial determiner elements consistently attach in just that way.”

(Anderson 2005:112)

The solution offered here lies in understanding these two morphemes as fundamentally different types of elements. Anderson asks the following question:

“How are we to describe these determiner elements? The “pre-nominal” component is straightforward: it is a clitic, appearing at the left edge of the nominal. We could treat it as we have to this point, as an item appearing in a structural determiner position which is initial within the nominal. Alternatively, we could describe it as a left-edge special clitic, but the choice between these two alternatives is not clear...”

(Anderson 2005:103-104)

Under the current proposal, the choice between these alternatives is clear and has observable consequences. The answer is that the prepositional determiner is a syntactic head merged by *Concatenate* while the 2P clitics are NP features merged by *Adjoin*.<sup>11</sup> This analysis is in fact strongly supported by the content of the determiners themselves. Observe the various K<sup>w</sup>ak<sup>w</sup>ala (unpossessed) determiner elements in Table 4.1.

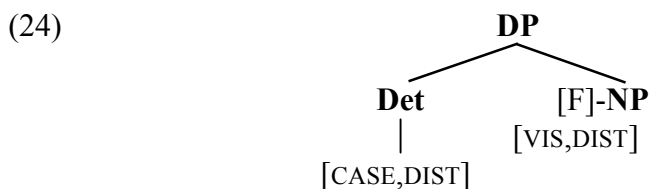
---

<sup>11</sup> That the prepositional determiner in K<sup>w</sup>ak<sup>w</sup>ala does not require to be positioned by special principles is also noted by Embick & Noyer (1999:292-3). In their framework (see chap. 5 below), this excludes the preposition from being treated as a true “ditropic” clitic (i.e., a clitic with contradictory phonological and syntactic dependencies). Rather, these morphemes are simply termed “left-leaners” without discussing by what principles the preposition is parsed with preceding material. They identify ditropic clitics as only those elements which have undergone “Local Dislocation” and are thus not in their expected position from the point of view of the syntax. Such elements have “structural hosts”, unlike elements positioned by regular syntax. This appears to be a circular definition of clitic and without clear principles for prosodic parsing, claims about the role of phonological attachment remain unclear.

**Table 4.1.** K<sup>w</sup>ak<sup>w'</sup>ala DP demonstrative clitics (unpossessed) (Anderson 2005:102)

		1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
<b>Pre-nominal</b>	<b>Subj</b>	-ga(da)	-ux(da)	-i(da)
	<b>Obj</b>	-x'ga(da)	-x <sup>w</sup> a, -x'ux'(da)	-x'(a)
	<b>Inst</b>	-sga(da)	-sa, -sux(da)	-s(a)
<b>Post-nominal</b>	<i>visible</i>	-k	-ix'	-∅
	<i>invisible</i>	-ga	-aq <sup>?</sup>	-a/i

Concurring with the claim that prenominal determiners are syntactic heads, it is only these elements which contain features relevant to external syntax, i.e., case features. The post-nominal elements, on the other hand, only indicate distal and visibility distinctions and are compatible with any of the three cases. Distal and visibility distinctions are (presumably) not relevant for sentential syntax and certainly cannot be understood to define the label of the larger phrase, unlike the case markers. Furthermore, the distal distinction in the 2P determiners appears to be a type of agreement, as it is also signaled by the head determiner. The only feature signaled uniquely by 2P determiners is visibility. The relevant (but partial) “underlying” structure for K<sup>w</sup>ak<sup>w'</sup>ala DPs is shown in (24). The features of 2P clitics, symbolized here by [F], are adjoined to the left edge of the NP while the head of the DP is concatenated to an XP complement. Crucially, the non-head features are outside the structure proper.



The relevant structure building operations here are thus *Concatenate* {D, NP} and *Concatenate* {V, DP}, which result in a head-initial determiner phrase and VP.

The noun phrase deictic features are added by *Adjoin* {[VIS,DIST], NP}. Crucially, the deictic features are associated with and hence adjoined to NP and not DP. Now, recall the PCP from (12) above, repeated here as (25):

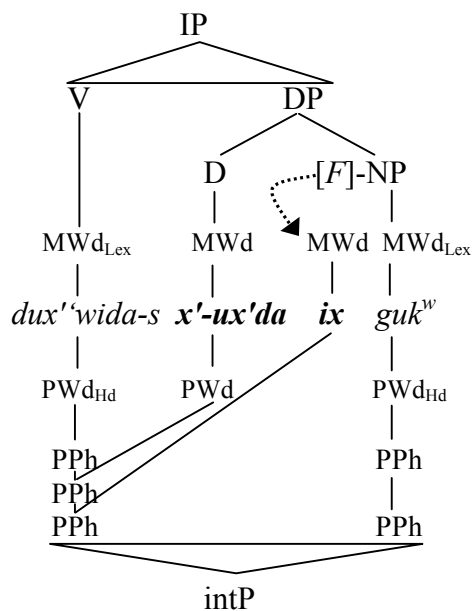
- (25) Phase Correspondence Principle (PCP):  
 Features are spelled out within the phase which they are attached

If DP is a strong phase as has been suggested by Chomsky (2001) and argued for explicitly by Gutierrez-Bravo (2001) *inter alia*, then merging of D will trigger Spell-Out of its complement, which we label here as NP. The same constraint against “cross-phase clisis” seen earlier should then apply in the DP domain. Enclisis of the deictic clitics to the determiner can now be ruled out if they are the spell-out of features attached to NP. With this in mind, let us now examine the evaluation of the K<sup>w</sup>ak<sup>w</sup>’ala sentence in (23) with the structure in (24). The three most likely candidates for the output of (24) are shown in (26)-(28). Constituents created by *Concatenate* are represented by standard binary branching phrase structure. The relation between the surface position of adjoined morphemes and their underlying position is indicated by the dotted arrow. As above, the relevant violation profile accompanies each structure.

In all three candidates considered below, the concatenated case marker encliticizes to the preceding verb, violating ALIGN (XP;PPh) in order to satisfy \*WEAKSTART (PPh). In (26), the adjoined NP features are spelled out faithfully at the left edge of NP, satisfying ALIGN (Morph;[F]) and the deictic clitic is parsed with the preceding prosodic phrase. This prosodic parse satisfies \*WEAKSTART but violates the PCP, a potentially inviolable (architectural) constraint. In (27), the order of elements is the same but the deictic clitic is parsed with the following prosodic phrase, satisfying the PCP at the cost of violating \*WEAKSTART. In (28), the grammatical order, the adjoined NP features are spelled out following the noun, violating ALIGN (Morph;[F])

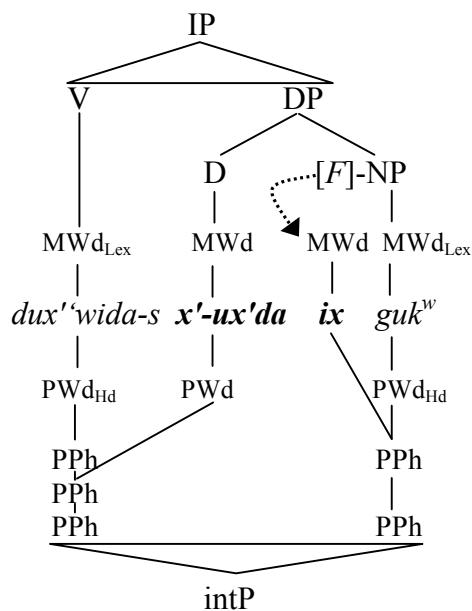
once because the morpheme is separated from the adjunction site, but satisfying both the PCP and \*WEAKSTART.

(26)



violation profile	
ALIGN (Morph;[F])	✓
PCP	✗
*WEAKSTART (PPh)	✓
ALIGN (XP;PPh)	✗

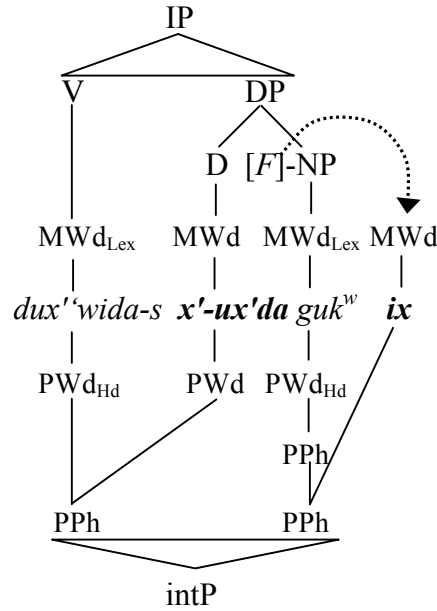
(27)



violation profile	
ALIGN (Morph;[F])	✓
PCP	✓
*WEAKSTART (PPh)	✗
ALIGN (XP;PPh)	✗✗



(28)



violation profile	
ALIGN (Morph;[F])	✗
<b>PCP</b>	✓
*WEAKSTART (PPh)	✓
ALIGN (XP;PPh)	✗

If the PCP is inviolable, we can immediately rule out (26) in favor of (27) and (28). The choice between (27) and (28) is left to the relative ranking of ALIGN (XP=PPh), \*WEAKSTART (PPh) and ALIGN (Morph;[F]). In the next subsection we see what the typological predictions of the above constraints are in all of the possible rankings.

#### 4.5 The factorial typology of adjoined clitics

The factorial typology of parsing interactions between concatenated and adjoined clitics can be derived solely with the constraints employed above. Tableaux 4.4-4.7 show how the different options are derived. We exclude here candidates which violate the PCP by parsing the adjoined clitic separately from the phrase to which it is adjoined to. For expository ease, we take a situation in which all elements, functional and lexical, satisfy prosodic wordhood. We furthermore assume that  $MWd_{Lex}=PWd_{Hd}$  is highly ranked so that functional elements may be parsed as prosodic words but only lexical elements will be parsed as prosodic word heads (abbreviated below as  $\omega$  and  $\omega_{Hd}$ ).  $X^0$  represents a concatenated syntactic head with YP (containing a single lexical

word) as its complement. *Cl* represents an adjoined clitic with its adjunction site as indicated in the input.

Tableau 4.4 offers two winning candidates, (a) and (c), which differ only in the amount of prosodic structure created. When YP represents the Spell-Out domain of a phase, candidate (a) can be eliminated on the grounds that each domain must project its own prosodic phrase. The adjoined clitic surfaces in its adjunction site because of the high-ranked ALIGN (Morph;[F]) and  $X^0$  is parsed with the following complement and not preceding material because ALIGN (XP;PPh) outranks \*WEAKSTART (PPh).

<b>Input:</b> $XP[X^0 \varphi_{-YP}[Y_{Lex}]]$	ALIGN (Morph;[F])	ALIGN (XP;PPh)	*WEAKSTART (PPh)
a. $\begin{matrix} \text{☞} \dots ] XP[X^0 cl_{YP}[Y_{Lex}]] \\ PPh[\omega \quad \omega \quad \omega_{Hd}] \end{matrix}$		*	**
b. $\begin{matrix} \dots ] XP[X^0 \quad cl_{YP}[Y_{Lex}]] \\ \dots \omega ] PPh[\omega \quad \omega_{Hd}] \end{matrix}$		**!	*
c. $\begin{matrix} \text{☞} \dots ] XP[X^0 \quad cl_{YP}[Y_{Lex}]] \\ PPh[\omega \quad PPh[\omega \quad \omega_{Hd}]] \end{matrix}$		*	**
d. $\begin{matrix} \dots ] XP[X^0 \quad YP[Y_{Lex} cl]] \\ PPh[\omega \quad PPh[\omega_{Hd} \omega]] \end{matrix}$	*!		*
e. $\begin{matrix} \dots ] XP[X^0 \quad YP[Y_{Lex} cl]] \\ \dots \omega ] PPh[\omega_{Hd} \omega] \end{matrix}$	*!	*	

**Tableau 4.4.** ALIGN (morph) >> ALIGN (XP,PPh) >> \*WEAKSTART

In Tableau 4.5, the winning candidate, (d), positions the adjoined clitic after the complement in 2P because the PCP requires parsing of the clitic with its host while the high ranked ALIGN (XP;PPh) requires the left edge of the prosodic phrase be aligned to the syntactic phrase. The optimal solution is thus displacement. The latter constraint also results in the proclitic parsing of the syntactic head  $X^0$ .

<b>Input:</b> XP[X <sup>0</sup> φ-YP[Y <sub>Lex</sub> ]]	ALIGN (XP;PPh)	ALIGN (Morph;[F])	*WEAKSTART (PPh)
a. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[ω ω ω <sub>Hd</sub> ]	*!		**
b. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] ...ω] PPh[ω ω <sub>Hd</sub> ]	*!*		*
c. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[ω PPh[ω ω <sub>Hd</sub> ]]	*!		*
d. ☞ ...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] PPh[ω PPh[ω <sub>Hd</sub> ω]]		*	*
e. ...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] ...ω] PPh[ω <sub>Hd</sub> ω]	*!	*	

**Tableau 4.5.** ALIGN (XP,PPh) >> ALIGN (morph) >> \*WEAKSTART

In Tableau 4.6, the winning candidate, (e), parses X<sup>0</sup> with preceding material and places the adjoined clitic in 2P within its domain. Together, the adjoined clitic and the lexical YP which constitutes its domain are parsed into a single prosodic phrase.

<b>Input:</b> XP[X <sup>0</sup> φ-YP[Y <sub>Lex</sub> ]]	*WEAKSTART (PPh)	ALIGN (XP;PPh)	ALIGN (Morph;[F])
a. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[ω ω ω <sub>Hd</sub> ]	*!*	*	
b. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] ...ω] PPh[ω ω <sub>Hd</sub> ]	*!	**	
c. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[ω PPh[ω ω <sub>Hd</sub> ]]	*!*	*	
d. ...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] PPh[ω PPh[ω <sub>Hd</sub> ω]]	*!		*
e. ☞ ...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] ...ω] PPh[ω <sub>Hd</sub> ω]		*	*

**Tableau 4.6.** \*WEAKSTART >> ALIGN (XP,PPh) >> ALIGN (morph)

Finally, in Tableau 4.7, the winning candidate in (b) parses X<sup>0</sup> with preceding material but places the adjoined clitic in domain initial position and parses it together with the following material in a single prosodic phrase. This is derived by the intermediate ranking of \*WEAKSTART where it is dominated by ALIGN (XP;PPh).

With this ranking, \*WEAKSTART exerts an influence on the parsing of concatenated head clitics but not the position of adjoined clitics.<sup>12</sup>

<b>Input:</b> XP[X <sup>0</sup> φ-YP[Y <sub>Lex</sub> ]]	ALIGN (Morph;[F])	*WEAKSTART (PPh)	ALIGN (XP;PPh)
a. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[ω ω ω <sub>Hd</sub> ]		**!	*
b. ☞ ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] ...ω] PPh[ω ω <sub>Hd</sub> ]		*	**
c. ...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[ω PPh[ω ω <sub>Hd</sub> ]]		**!	*
d. ...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] PPh[ω PPh[ω <sub>Hd</sub> ω]]	*!	*	
e. ...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] ...ω] PPh[c]	*!		*

**Tableau 4.7.** ALIGN (morph) >> \*WEAKSTART >> ALIGN (XP,PPh)

Table 4.2 shows a summary of the possibilities and the rankings which derive them. Note that not all the combinatorial possibilities for the ranking of the three constraints are given as the high ranking of ALIGN (XP;PPh) or \*WEAKSTART (PPh) renders the relative ranking of the other two constraints moot.

**Table 4.2.** Factorial typology of concatenated and adjoined clitic interactions

<b>Output</b>	<b>Ranking</b>
a.i. [...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] [...] PPh[ω ω ω <sub>Hd</sub> ]	ADJOIN-L (F,YP) >> ALIGN (XP;PPh) >> *WS
ii. [...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[ω [ω ω <sub>Hd</sub> ]]	
b. [...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] PPh[ω [ω <sub>Hd</sub> ω]]	ALIGN (XP;PPh) >> *WS, ADJOIN-L (F,YP)
c. [...] XP[X <sup>0</sup> YP[Y <sub>Lex</sub> cl]] PPh[... ω] [ω <sub>Hd</sub> ω]	*WS >> ALIGN (XP;PPh), ADJOIN-L (F,YP)
d. [...] XP[X <sup>0</sup> cl <sub>YP</sub> [Y <sub>Lex</sub> ]] PPh[... ω] [ω ω <sub>Hd</sub> ]	ADJOIN-L (F,YP) >> *WS >> ALIGN (XP;PPh)

<sup>12</sup> Needless to say, in order to derive the typologically marked prosodic phrase initial position for adjoined clitics in Tableaux 4.5 and 4.2, ALIGN (PPh<sup>max</sup>, XP) must also be ranked low.

The K<sup>w</sup>ak<sup>w</sup>ala pattern discussed by Anderson exemplifies the ranking in (c). We can thus offer a principled explanation for the co-presence of enclitic prepositions and 2P clitics in the same grammar. More common 2P clitic systems such as that of Tagalog, with proclitic prepositions, correspond to the ranking in (b). Because we do not have very complete information on languages which instantiate 1P clisis (e.g., Tondano) we cannot yet offer a definitive exemplar for the rankings in (a) or (d), although because both of these languages have clitic adpositions and case markers, we can be sure that at least one of these patterns is attested.

In the following sections we take a look at other predictions of the *Adjoin* and *Concatenate* model in similar domains.

#### **4.6 Further consequences**

The distinction proposed here between Merge Feature (*Adjoin*) and Merge Terminal (*Concatenate*) makes several other predictions that are worth exploring here. In this section, we take a brief look at head-adjacent clisis, differences between clitic and free pronominals and morphological feature cooccurrence constraints.

##### *4.6.1 Verb adjacent clitics and affixes*

Assuming a degree of parallelism between the construction of sentential structure and word structure (Baker 1985, Lieber 1992, Halle & Marantz 1993 *inter alia*), the simplest hypothesis is that words may also be built by the merger of syntactic heads via *Concatenate* as well as merger of features by *Adjoin*. If this is the case we expect to find parallels between the phrasal typology explored in chap. 3 and word level morphological phenomena. Two such potential parallels are proposed here.

Verb adjacent clitics in Medieval Romance languages and modern Bulgarian possess a default alignment to the verb which is apparently subverted in particular

prosodic contexts. Specifically, these verb adjacent clitics are proclitic on the verb unless this places them in initial position of a relevant prosodic or syntactic domain in which case they encliticize. The relevant domain has been argued at length by Fontana (1996) and Wanner (1987, 1996) to be prosodic in nature. This positioning pattern, which can be summarized as “proclitic unless initial”, is commonly referred to as Tobler-Mussafia effects, after their first analysts (Tobler 1889, Mussafia 1898). Tobler-Mussafia effects can be understood here to result from \*WEAKSTART taking effect on verb adjoined, rather than clause adjoined clitics. Similar to other OT approaches (cf. Legendre 2000a, 2000b), this can be derived by alignment of a morpheme adjoined to a verbal or inflectional head.

Unlike previous OT analyses, the added explanatory power of the present theory allows us to predict the different behaviors two types of Romance clitics which are phonologically and etymologically identical but functionally very different. The strong phonological resemblance between object clitics and determiners in Romance has led various researchers (Cardinaletti & Roberts 1991, Uriagereka 1995, Martins 1995) to treat them both as D elements with null or *pro* complements. Fischer (2002) notes a problem for this analysis in that the phonological constraint which bans clitics in first position must apply only to object clitics but not to their homophonous determiner counterparts, which are unswervingly proclitic. Compare the proclitic position of /la/ and /lo/ in their determiner function in (29) and (30) with the enclitic position of the pronominal in the second conjunct in (30).

(29) ...e **la**=sant crestià provava al juheu que... *Old Catalan*  
 CONJ DET=saint Christ proved.3s to.DET Jew that  
 ‘...and the saint Christ proved to the Jew that...’ (Fischer 2002:134)

(30) ...e presentà’s denant **lo**=rey e saludà=**lo** molt altament  
 CONJ present.3s before DET=king CONJ greet=3s very highly  
 ‘...and he appeared before the king and greeted him warmly...’ (ibid.)

Fischer (2002:134) sums up the problem: “In this case the constraint would not only have to read the syntactic label, but it would also have to interpret the semantics and/or to analyze whether or not a complement is taken by the  $D^0$ .” The present theory allows us to predict the above facts elegantly while maintaining the desirable generalization that the D head is essentially the same element in both cases.<sup>13</sup> The morphemes /lo/ and /la/ head a syntactic phrase (the DP) in their determiner function but only express argument features in their object function. Accordingly, as phrasal heads, they must be introduced into the syntax via *Concatenate*, but as argument features they are introduced by *Adjoin*. When introduced by *Adjoin*, they are subject to misalignment due to \*WEAKSTART. When introduced by *Concatenate*, on the other hand, only syntactic D-stranding movement is able to separate the head from its complement, a movement which is unavailable for purely syntactic reasons in Romance.<sup>14</sup> The difference in underlying representations is illustrated in (31).

---

<sup>13</sup> Thanks to Satoshi Tomioka for discussion on this point, especially regarding the value of this generalization.

<sup>14</sup> Fischer (2002, 2003) actually argues against a Tobler-Mussafia interpretation of Catalan clitics, citing examples of clitic-verb order in initial position and verb-clitic order in seemingly non-initial position. Her solution is semantically grounded in that verb-clitic order is uniformly derived by (semantically triggered) movement to the head of PolarityP (Laka 1990). Nonetheless the Tobler-Mussafia pattern seems to hold in a significant majority of cases in her corpus study of Old Catalan and there exist apparent exceptions to the Tobler-Mussafia law in every language for which it has been discussed. A serious problem for statistical counts of clitic position in historical texts is that not all topicalizations are indicated orthographically (e.g., by a comma). This could often give the impression that verb-clitic order obtained despite the presence of a preceding host when in fact the post-verbal order results from the topic phrasing of the preceding material. The converse possibility of clitic-verb in initial position could indicate exaptation by which the order of elements itself indicated conjunction, even without the overt presence of the conjunction head itself. This is similar to what is described for Yami (Rau & Dong 2006) and also for Pamona (van den Berg 2002, Mead 2002) where proclisis had come to indicate conjunction whereas it was previously only a concomitant of the preceding conjugation head which functioned as a clitic host. Finally, it should be mentioned that, while the “syntacticization” of clitic placement from old to modern Romance is incontrovertible, Fischer does not offer a semantic explanation for a sufficient number of textual cases to cast serious doubt on the role of the phonological component in the older texts.



For the sake of concreteness, we can offer the following sketch for the derivation of the adjoined object clitics. As illustrated in (31)b, I take Catalan clitics to be the spell-out of object features adjoined to the left edge of the V.<sup>15</sup> I further take \*WEAKSTART (iP) to be highly ranked so that the left edge of intonational phrases prefer to be aligned to strong elements. As Fischer (2002) discusses, enclisis is typically found on sentence initial verbs but also appears on verbs following certain other constituents within the clause. From Fischer’s examples, most of these constituents can easily be reanalyzed as topics, and there could have thus been an intended intonational phrase boundary which disallowed proclisis. While I take conjunction to have optionally induced an intonational phrase boundary following it, there was no such option for negation, which was obligatorily included in the following intonational phrase.<sup>16</sup> As a result, the possibility of proclisis after conjunction depended on the inclusion of conjunction in the following intonational phrase. To simplify matters, we evaluate these possibilities independently. Tableau 4.8 shows the evaluation when the conjunction is not included in the following intonational phrase and Tableau 4.9 shows the evaluation when it is included. Enclisis obtains in the first case and proclisis in the latter.

<sup>15</sup> More precisely, I<sup>0</sup>, as finite auxiliaries host clitics when present.

<sup>16</sup> Mériz (1978:305-306, via Wanner 1987:157) discusses the role of sentential vs. VP conjunction on clitic position in Old Provençal noting that enclisis is more common with sentential coordination while proclisis is more common with VP coordination. A similar observation of made by Fontana (1996:75) for Medieval Spanish.



Input: <i>e</i> [φ]- <i>saluda</i> CONJ [3S.MSC]-greet	*WEAKSTART (iP)	ALIGN (Morph;[F])
a. $\leftarrow$ <i>e</i> <i>saludà lo</i> cl <sub>iP</sub> [ <sub>PPh</sub> [PWd=cl]]		*
b. <i>e</i> <i>lo saluda</i> cl <sub>iP</sub> [ <sub>PPh</sub> [cl=PWd]]	*!	

**Tableau 4.8.** Tobler-Mussafia enclitics following a conjunction

Input: <i>e</i> [φ]- <i>saluda</i> CONJ [3S.MSC]-greet	*WEAKSTART (iP)	ALIGN (Morph;[F])
a. <i>e</i> <i>saluda lo</i> iP[cl <sub>PPh</sub> [PWd=cl]]		*!
b. $\leftarrow$ <i>e</i> <i>lo saluda</i> iP[cl <sub>PPh</sub> [cl=PWd]]		

**Tableau 4.9.** Tobler-Mussafia proclitics following a conjunction

When negation precedes the verb, there is never an option for enclisis. This makes sense, as there is also no option for topicalization of negation, which is a functional head taking a following clausal complement. The evaluation of proclisis with negation is shown in Tableau 4.10.<sup>17</sup>

Input: <i>e no</i> [φ]- <i>meravell</i> CONJ NEG [1S]-surprise	*WEAKSTART (iP)	ALIGN (Morph;[F])
a. <i>e no meravell me</i> cl <sub>iP</sub> [ <sub>PPh</sub> [PWd PWd=cl]]		*!
b. $\leftarrow$ <i>e no em meravell</i> cl <sub>iP</sub> [ <sub>PPh</sub> [PWd cl=PWd]]		

**Tableau 4.10.** Tobler-Mussafia proclitics following Neg

In line with the interpretation of Tobler-Mussafia clitics as merged features we find that they are exclusively made up of pronominals in both Romance and Bulgarian. Because pronominals consist solely of phi-features and do not project their syntactic label they are fully compatible with *Adjoin*. Note that similar cases of

<sup>17</sup> Evidence that proclisis was sensitive to preceding material can be seen in the allomorphy between the 1<sup>st</sup> sg. object clitic /me/ which (counter-intuitively) appears as /em/ after a vowel (Fischer 2003:277).

phonologically conditioned morpheme positioning have been observed on the word level, as well, two prominent cases being that of Afar (Fulmer 1990, Billings 2004), shown in (32), and Huave (Noyer 1994, Billings 2004, Kim 2008), shown in (33).<sup>18</sup>

(32) a. *t*-ubl-é                      b.    suk-*t*-é                                      *Afar*  
           3.FEM/2-see-PRF            have-3.FEM/2-PRF  
           ‘she/you saw’                ‘she/you had’                                      (Fulmer 1990: 190)

(33) a. *t*-awit’                        b.    wit’i-*t*                                      *Huave*  
           PAST-raise.TR                rise.INTR-PAST  
           ‘[s]he raised [it] up’        ‘[s]he rose up’                                      (Noyer 1994: 71)

In both cases, the morpheme in question is coincidentally of the shape /t/. In Afar, *t* is an agreement marker signaling either third person singular feminine or second person. It is generally a suffix except when attaching to a (non-low) vowel initial stem in which case it is a prefix. The Huave affix, first described by Noyer (1994), is similarly mobile, appearing prefixally in certain contexts with vowel initial stems. Of relevance to the present discussion is the fact that, although not many cases of mobile affixes have been reported in the literature, all described cases are thoroughly inflectional (agreement and tense) rather than derivational. This distribution is expected on the present approach since only inflectional morphology is compatible with *Adjoin* and only affixes attached by *Adjoin* could have variable polarity.<sup>19</sup>

One prediction which is more difficult to tease out regards the applicability of the PCP to Tobler-Mussafia clitics. Old Spanish pronominal clitics appear to be 2P

<sup>18</sup> See Paster (2006), however, for arguments that these do not represent phonologically motivated affixation but rather instantiate morphological subcategorization.

<sup>19</sup> In fact, this prediction is only met half way since there is ample attestation of derivational infixes. Infixation must thus not be completely parallel to 2P clisis, as only the latter is restricted to an inflectional/adverbial function. It was argued however that what prevents syntactic heads from entering 2P is the lack of legitimate syntactic movement which could regularly derive this positioning. The same argument must clearly be modified for infixation as we do not yet have a clear notion of what constitutes illegitimate “movement” within the word.

clitics which take some CP layer as their domain and can follow  $C^0$  hosts such as *que* and *si*. However, following the development of these pronominals into verb adjacent Tobler-Mussafia clitics, there exists a pattern of avoiding complementizer hosts. As discussed by Fontana (1996:75), Mériz (1978:305-306, via Wanner 1987:157) and others, clisis patterns differently in sentential and VP conjunction in several old Romance languages (e.g., Medieval Spanish, Old Provençal). The preferred pattern is shown in (34), where proclisis is preferred with VP conjuncts but enclisis is preferred with sentential conjuncts.

- (34) a.            CONJ-*cl*-(XP)-V[+finite]                    *VP conjuncts*  
       b.            CONJ-V[+finite]-*cl*                         *Sentential conjuncts*

Although complementizers persist as potential hosts for verb adjacent Tobler-Mussafia clitics in the Medieval Romance, examples of clitic third with preceding complementizers are also common. We can tentatively attribute the inability of verb adjacent clitics to be hosted by complementizers as the result of the PCP as this would constitute cross-phase clisis. The possibility of clisis to complementizers on the other hand could be attributed to V to C movement for which there is ample independent evidence in Romance. If this approach is on the right track, then adjoined head-adjacent clitics display the same sensitivities to phase boundaries as do 2P clitics. It would thus follow that, concatenated clitics, whose anchoring to syntactic structure is more stable, may behave differently. Recall that the PCP, as formulated earlier, only refers to how adjoined features are spelled out. More generally, phonological interactions can take place across phase boundaries so long as the elements in question are terminal syntactic nodes. This is clearly necessary to account for the fact that there seems to be no restriction on the coalescence of heads across the various layers of DP

crosslinguistically – in particular across DP and NP – despite the fact that syntactic diagnostics suggest that DP appears to constitute a phase (Legate 2002, Svenonius 2004 *inter alia*). Extending the parallelism fully to the word level we thus expect derivational affixes which can only be added by *Concatenate* on the present theory, to be able to lean leftwards away from their hosts without triggering a violation of the PCP.<sup>20</sup> One such case is discussed at length by Dixon (1988) in his grammar of Boumaa Fijian. The prefix *i-* is a derivational morpheme which converts verbs to nouns, as can be seen from (35).

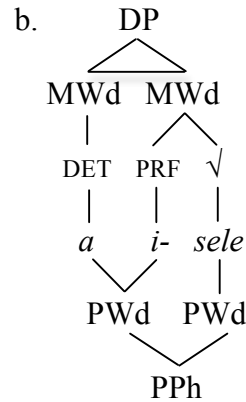
- |         |                                 |    |                                      |   |
|---------|---------------------------------|----|--------------------------------------|---|
| (35) a. | i-sele<br>NMLZ-cut<br>'a knife' | b. | i-talanoa<br>NMLZ-narrate<br>'story' | <i>Boumaa Fijian</i><br><br>(Dixon 1988:21) |
|---------|---------------------------------|----|--------------------------------------|---|

This prefix consistently forms a prosodic word with preceding articles when they are present as diagnosed by diphthong formation which only occurs within prosodic words and not across them. The phonetic output of (36)a is stated by Dixon to be [ajsele] with diphthongization, rather than the expected [aisele] with hiatus. The analysis suggested is shown in (36)b, where the determiner and prefix combine to form their own prosodic word.

---

<sup>20</sup> All statements regarding *Adjoin* and *Concatenate* on the word level are more tentative than phrase level generalizations. The phonological typology of derivational versus inflectional affixes is unclear and quite probably not nearly as categorical as that found with phrase level clitics. If derivational affixes can only be attached via word level *Concatenate* and infixation is truly akin to 2P within the word, then we would not expect derivational infixes. There are, however, several attestations of infixing nominalizers, e.g. in Katu (Horwood 2008) and Leti (van Englehoven 2004). In Tagalog, it is also likely that the actor voice infix <um> is best understood as derivational rather than inflectional. On the other hand, as discussed above, Rudin (et al 1999) have proposed for Bulgarian that 2P can operate within the X<sup>0</sup> level as well and it was suggested here that this pattern did not violate constraints on movement. Similarly, the word internal reordering found with infixation may be able to maintain underlying morphological bonds while satisfying phonological requirements.

(36) a. a i-sele  
 DET NMZ-cut  
 ‘the knife’



This is exactly parallel to the cases of prepositional enclitics discussed earlier. A weak (functional/vowel initial) prefix breaks with its morphological constituency and attaches to the left in order to align the PWd with a strong edge. We thus find both on the phrasal and word levels examples of concatenated morphemes breaking their morphosyntactic constituency to be parsed with material to the left. On the phrasal level this parsing was argued to be preferred on the basis of satisfying \*WEAKSTART (PPh). On the word level the same phenomenon could satisfy \*WEAKSTART (PWd), or a lower (syllable) level constraint (e.g. \*HIATUS or ONSET).

#### 4.6.2 2P versus free pronominals

We have yet to full address the difference in formal encoding between free pronominals and 2P pronominals. It was suggested in chap. 2, following Everett (1996) that the difference between agreement, 2P clitics and free pronouns is a function of where features are merged. More concretely, it is taken as a function of where an element is merged and how it is merged. Free pronouns are generally found to occupy argument positions and behave like full NPs in relevant respects. They are thus understood here to result from the merging of phi-features into argument position via *Concatenate*. 2P clitics, on the other hand, consistently display syntactic divergences from full NP arguments, both in their positioning and in other aspects of

their syntactic behavior. In this section I will briefly discuss the motivations for merging phi-features via *Adjoin* versus *Concatenate* and the relevance of featural cooccurrence restrictions.

In Slavic and Romance (in addition to Austronesian languages such as Chamorro, Cebuano, Tboli, and many others) we find that pragmatic factors can license non-clitic pronominals in canonical argument positions, as exemplified by the contrast Serbo-Croatian in (37)a, with a backgrounded pronominal object and b, with a focused pronominal object.

- (37) a. Da=**ti** dam knjigu? *Serbo-Croatian*  
 CONJ=2s give book  
 ‘Should I give you the book?’ (Browne 1974; Nevis 1988:69)
- b. Da dam knjigu **tebi**?  
 CONJ give book 2s  
 ‘Should I give YOU the book?’ (Browne 1974; Nevis 1988:69)

For many Philippine languages, this is impossible. In Tagalog and Ilokano, for instance, non-clitic forms can only appear in topic or predicate position and are generally banned from canonical argument positions. This is shown in (38) and (39).

- (38) a. Nag-reklámo=**ka** b. \*Nag-reklámo **ikaw** *Tagalog*  
 AV.BEG-complain=2s.NOM AV.BEG-complain 2s.NOM  
 ‘You complained.’
- (39) a. Nag-takder=**ak** b. \*Nag-takder **siak** *Ilokano*  
 AV.BEG-stand=1s.NOM AV.BEG-stand 1s.NOM  
 ‘I stood up.’

The Slavic/Romance situation is thus probably best described as pragmatic complementarity whereas the Tagalog/Ilokano situation in (38) represents true

syntactic complementarity.<sup>21</sup> The difference can be reduced to the ranking of a simple \*STRUCTURE constraint. Constraints of the \*STRUCTURE family enforce economy of pronunciation on several levels. Following Picallo (1994), I take the relevant level here to be referential features and for strong (free) pronouns to have more referential features than weak pronouns. As argued by Picallo, the richer referential structure of strong pronouns accounts for their lack of co-reference potential in Romance when compared to weak pronouns (see also Siewierska 2004, for discussion). Given  $\varphi_i$ , the referential agreement feature which distinguishes weak and strong pronominals, we can formulate our \*STRUCTURE constraint trivially as follows:

\*STRUCTURE ( $\varphi_i$ )  
 Violated by a pronominal in the output bearing the strong referential feature  $i$ .  
 (cf. Chomsky 1981, Cardinaletti & Starke 1994:48; Bresnan 2001)

The ranking in (40)a will then result either in the stripping of this feature in the mapping from input to output, or the expression of the pronominal as a predicate, i.e. the Tagalog and Ilokano state of affairs. The ranking in (40)b, will license free pronouns in argument position, as found in Romance and Slavic, among many other languages.

- (40) a.           \*STRUCTURE ( $\varphi_i$ ) >>> FAITH ( $\varphi_i$ )  
           b.           FAITH ( $\varphi_i$ ) >>> \*STRUCTURE ( $\varphi_i$ )

Although pragmatic motivations cannot license free pronouns in many Philippine languages, several languages employ free pronouns as a repair strategy for

---

<sup>21</sup> Note that the ban on free pronouns found in some Philippine languages only holds for unextracted arguments. The predicate and topic positions can only host free pronouns for both feature based and prosodic reasons. Predicates and topics are not compatible with reduced referential features because these elements are not backgrounded but are rather contrastive, in some sense. Prosodically, as well, the initial position occupied by predicates and topics in Philippine languages cannot host clitic pronouns given a highly ranked \*WEAKSTART constraint.

person case cooccurrence constraints. An extremely common person cooccurrence constraint in Philippine languages involves the combination of first person singular genitives with second person singular nominatives. In Tagalog this is resolved by suppletion. Rather than 1S.GEN *ko* and 2S.NOM *ka* surfacing together, they are both replaced by a single portmanteau morph *kita* 1S.GEN+1S.NOM. Other Philippine languages have other strategies for dealing with this constraint. Relevant to the 2P clitic - free pronoun distinction under discussion is the process of “disformation” (Billings & Kaufman 2004, Peng & Billings 2005, Kaufman to appear b), by which a pronominal is forced to appear in its free form. Disformation can be exemplified by Maranao (Kaufman to appear b). Maranao shows the Tagalog-type ban on free pronouns in argument position, as shown in (41). Example (41)a shows that the free pronoun *seka* 2S.NOM is licensed in topic position while (41)b and c show that, in the unmarked case, only the 2P clitic *ka* 2S.NOM is possible in post-predicate position.

- (41) a. **Seka** na s<om-iy>ŋ sa=iŋed *Maranao*  
 2S.NOM TOP <AV-PRF>go OBL=village  
 ‘You, went to the village.’
- b. S<om-iy>oŋ=ka sa=iŋed  
 <AV-PRF>go=2S.NOM OBL=village  
 ‘You went to the village.’
- c. S<om-iy>oŋ [\*seka] sa=iŋed [\*seka]  
 <AV-PRF>go 2S.NOM OBL=village

The same person case constraint found in Tagalog also exists in Maranao except here, instead of suppletion we find replacement of a clitic with a free pronoun. The ungrammatical (42)a is thus expressed as (42)b.

- (42) a. \*M<iy>a-ilay=**aken=ka** *Maranao*  
 <PRF>PV.NVL-see=1S.GEN =2S.NOM





(44) a. \*Há~hanáp-in=**ko=ka** *Tagalog*  
 INCM~search-PV=1S.GEN=2S.NOM

b. Há~hanáp-in=**kita**  
 INCM~search-PV=1S.GEN:2S.NOM  
 ‘I will look for you.’

(45) Ákin=**ka=ŋ** há~hanáp-in *Tagalog*  
 1S.GEN=2S.NOM=LNK INCM~search-PV  
 ‘I will look for you.’

(46) Ikaw aŋ=há~hanáp-in=**ko** *Tagalog*  
 2S.NOM NOM=INCM~search-PV=1S.GEN  
 ‘You’re the one I’ll look for.’

Person-case constraints are thus seen to be only triggered by adjoined features and not by concatenated arguments. We can now turn to the structural description for these constraints. Here we find interesting evidence from Agutaynen, which, as was briefly discussed in §2.4.1, is more liberal in not requiring clitics to cluster in a single position. The data in (47) suggests that the domain of person case effects is smaller than the clause. The relevant constraint in Agutaynen disforms a local agent in the presence of a local nominative patient. Thus, where we expect to find the genitive clitic *mo* 2S.GEN in (47)a, we instead find an oblique form *nio*. The datum of interest is shown in (47)b, where only the nominative attaches to negation, leaving the pronominal agent in the argument domain where it is presumably adjoined. In this case we find no disformation; the agent clitic surfaces as an expected genitive.

(47) a. Indi=**o** **nio** i-tabid *Agutaynen*  
 NEG=1S.NOM 2S.OBL IRR:PV-accompany  
 ‘Don’t include me!’ (Ruch & Quakenbush 2006:9)

b. Indi=**o** i-tabid=**mo**  
 NEG=1S.NOM IRR:PV-accompany=2S.GEN  
 ‘Don’t include me!’ (Ruch & Quakenbush 2006:9)

We can state the locality conditions easily in the current framework as adjacency at the adjunction site. Unlike Tagalog, Agutaynen allows adjunction of the agentive features directly to the theta-domain, understood here as  $\nu\text{P}$ . As at least certain types of  $\nu\text{P}$  are also considered to constitute phases, the PCP applies to this smaller domain, disallowing adjoined clitics from finding an external host, that is, prohibiting cross-phasal clisis.<sup>22</sup> Consequently, a clitic whose features are adjoined to  $\nu\text{P}$  will be placed after the first legitimate host in  $\nu\text{P}$ , rather than remaining at its left edge. The variation in positioning and morphological interaction between the clitics is thus predicted by the two adjunction structures in (48)a and b (assuming non-crucially that Neg is a TP adjunct), corresponding to (47)a and b, respectively.

- (48) a.  $[\varphi_{\text{NOM}}]-[\varphi_{\text{GEN}}]-\text{TP}[\text{Neg}_{\text{TP}}[\nu\text{P}[\text{V}]]]$   
 b.  $[\varphi_{\text{NOM}}]-\text{TP}[\text{Neg}_{\text{TP}}[[\text{F}_{\text{GEN}}]-\nu\text{P}[\text{V}]]]$

#### 4.7 Conclusion

We have seen in this chapter several pieces of evidence which further support the distinction between Merge Feature and Merge Terminal. An underlying theme of the chapter has been to demonstrate the benefits of being housed under a syntactic node. In the first half of the chapter it was shown how prosodic parsing on the phrase level includes bona fide syntactic heads while excluding adjoined features. This offers a strong motivation for adjoined features to evacuate their edge position. The same pressures effect bona fide syntactic heads but are satisfied by them in a different way, namely leaning away from their complement in-situ. In the second half of the chapter we observed how terminal nodes also shield their content from purely morphological constraints. Feature cooccurrence constraints were seen across several Philippine

---

<sup>22</sup> See Legate (2003) and den Dikken (2006, 2007) for two different views on the relevant type of  $\nu\text{P}$ .

languages to effect pronominal clitics while ignoring free pronouns. This comports well with a view of pronominal clitics entering the derivation via Merge Feature and free pronouns entering via Merge Terminal. When locality constraints could be observed directly it was furthermore seen that the proper domain for these constraints (in the languages investigated) was simple adjacency on the edge of a host.

The second contribution of this chapter has been to introduce the PCP (Phase Correspondence Principle). While it is not ruled out here that further study may show this constraint to result directly from the grammatical architecture (cf. Fox & Pesetsky 2004), as a preliminary descriptive principle it is able to account for the ungrammaticality of several clitic attachment possibilities on the phrase level (as seen for Tagalog and K<sup>w</sup>ak<sup>w</sup>ala) and possibly on the word level, as well (as seen in certain tendencies of Tobler-Mussafia effects).<sup>23</sup>

In the next chapter we undertake detailed investigation of Tagalog phrase structure and its consequences for clitic placement. There we will focus on the evidence for impenetrability phenomena, i.e., those cases in which an apparent potential host is barred from hosting clitics due to its position within the larger syntactic construction.

---

<sup>23</sup> Another obvious consequence of the difference between Merge Feature and Merge Terminal is the licensing of clitic climbing. In an extremely diverse range of languages, pronominal clitics are able to raise to a higher predicate from an embedded position in a way that full NP arguments cannot. It is a prediction of the present theory that adjoined clitics are not be subject to the same strictures on syntactic movement that syntactic nodes are and thus clitic climbing is derived easily. This is discussed in §5.4.

## CHAPTER 5: TAGALOG CLITIC SYNTAX AND THE CLITIC VISIBILITY CONDITION

### 5.1 Introduction

In the previous chapter a theory of feature adjunction was presented in order to account for certain facts of 2P clitics and their relation to other edgebound elements. In this chapter we now turn to the third question outlined in the introduction, namely, what constitutes the first element for clitic placement and what accounts for the impenetrability of certain constituents to clitic placement? As will be shown here, there are a number of contexts which allow 2P clitics to follow complex constituents in Tagalog (as discussed by S&O 1972, Sityar 1989, Kroeger 1993, Billings & Konopasky 2003).<sup>1</sup> The leading idea here in approaching impenetrability phenomena is that a syntactic filter operates on surface structure even if the ordering mechanisms responsible for clitic placement are not part of the general syntax. While not occupying terminal nodes within the phrase structure, 2P clitics are still subject to a locality relation with their predicates. This resolves an interesting paradox encountered with Austronesian 2P clitics. On the one hand, there is no syntactic position which can be associated with 2P clitics and no standard movement operations which can create the non-constituent hosts upon which 2P clitics lean, while on the other hand, there is a class of syntactic constituents which clitics cannot interrupt.

This chapter shows the behavior of 2P clitics within a wide range of syntactic environments. In addition to offering an improved theory of clitic placement for Tagalog, this chapter aims to catalogue the relevant data for a theory of 2P syntax in

---

<sup>1</sup> One prominent claim in the literature is that of Sityar (1989), who argues that Tagalog pronominal clitic placement is asymmetric across lexical categories. This claim was later taken up by Kroeger (1993) as evidence for the non-configurational nature of Tagalog and its lack of a VP and by Carnie (2000), who claims a related asymmetry in Irish. On further investigation, the empirical claim does not hold water, as discussed below.

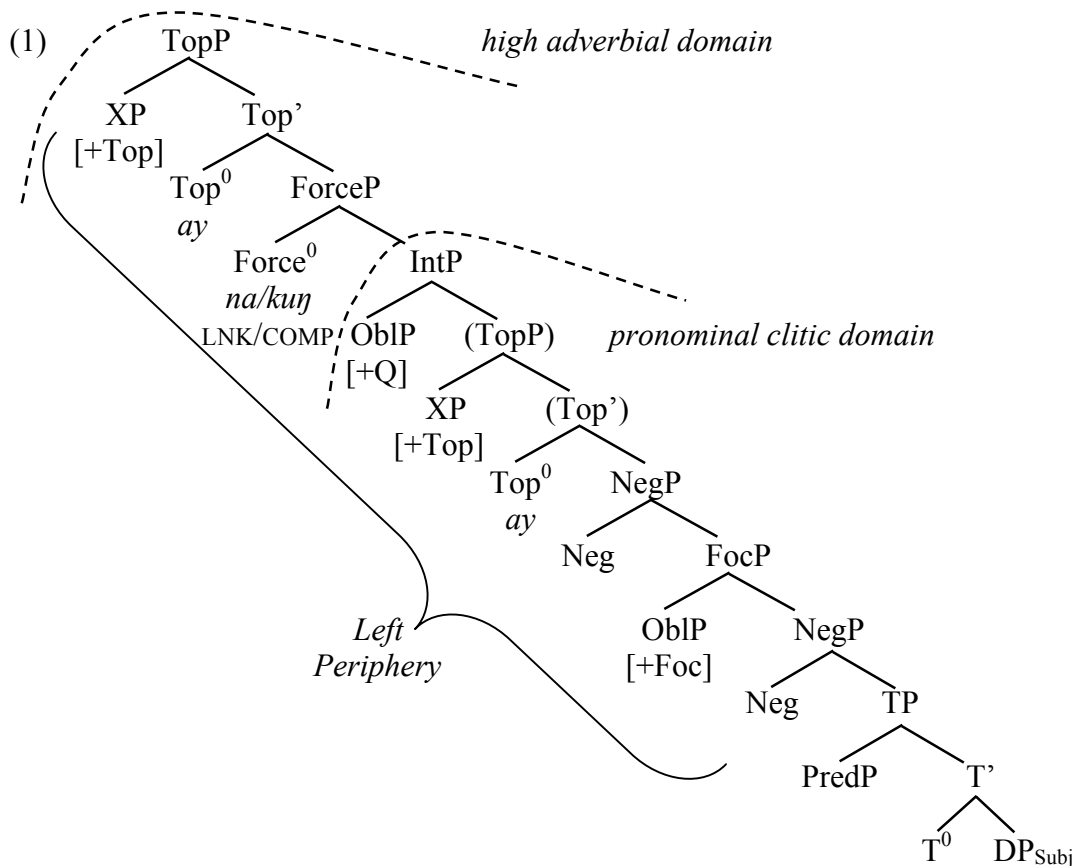
Tagalog and is thus particularly rich in data. As in the previous chapters, the examples are culled from texts wherever possible. Attested examples which appeared marginal were further checked with native speakers.

## **5.2 Clitic positioning within the clause**

In this section we will investigate the domains of clitic placement in Tagalog in relation to the left peripheral elements. This will require a closer look at the general clause structure of the language, in particular the loci of focus movement and topicalization.

### *5.2.1 The architecture of the left periphery*

The left periphery of the Tagalog clause, as proposed here, is shown in (1). It is bisected by the clitic domain, which reaches up to IntP, the phrase hosting interrogatives. The unorthodox configuration of TP is defended in Kaufman 2009 and will not be discussed here at any length.



In (1), we replace the generalized cover term CP with more specific projections which dominate TP.<sup>2</sup> Dominating TP we find a Focus Phrase sandwiched between two negation phrases. While it could easily be the case that negation is best considered an adjunct, for the sake of explicitness we include it here as one of the functional projections. The Focus Phrase houses an oblique phrase focus in its specifier. The example in (2) shows how the two negation elements can cooccur with the Focus Phrase.

- (2) NegP[Hindi FocP[sa=opisína NegP[hindi TP[k<in>á~ka-úsap si=Mario]]]]  
 NEG OBL=office NEG <BEG>~INCM-co-speak P.NOM=M.  
 ‘It’s not at the office that Mario’s not spoken to.’

<sup>2</sup> We do not consider here such phrases as FinP (Rizzi 1997, Mercado 2002) and others which may also find a place pending further analysis.

Immediately dominating the higher NegP is TopP, whose specifier hosts internal topics and is headed by the topic marker *ay*.<sup>3</sup> Dominating TopP we find IntP which hosts oblique interrogative elements in its specifier.

The possibility of cooccurring oblique interrogatives with oblique foci, as seen earlier in (164), is made possible by the fact that IntP exists alongside FocP and both can simultaneously be filled. IntP also demarcates the edge of the pronominal clitic domain above which only external topics and complementizers appear in ForceP.

ForceP requires some additional comment as it involves certain complications. First, note that the complementizer is positioned higher than usual. Instead of the more familiar configuration  $_{CP}[wh- C'[_{COMP}$ , e.g., as evidenced directly by Middle English (3), among others, we find that, in Tagalog, the complementizer appears higher than interrogative elements, as seen in (4).

- (3) men shal wel knowe **who that** I am *Middle English*  
 ‘One shall well know who I am’  
 (Caxton, 1485, R 67, in Lightfoot 1979: 322)

- (4) Hindi=niya s<in>ábi-Ø  $_{ForceP}$ [**kun**  $_{IntP}$ [saan=siya Ø-pú~punta]]  
 NEG=3S.GEN <BEG>say-PV COMP where=3S.NOM AV-INCM~go  
 ‘He didn’t say where he’ll go.’

Because, the complementizer *kun* only introduces conditionals and non-interrogative questions, we can safely identify it as belonging to ForceP, a projection which hosts illocutionary force related elements (Rizzi 1997:283).

Also interesting is the complementary distribution between external topics and complementizers. In (5), an external topic, in [Spec,ForceP], precedes an interrogative clause (which in turn contains a focus fronted oblique).

---

<sup>3</sup> See Aissen 1992 for the notion of internal and external topics and their manifestation in Mayan.



- (5) ForceP[[aŋ=maŋa=táo] ay IntP[kailan FocP[sa=ákin RP[Ø-lá~lápít]  
 NOM=PL=person TOP when OBL=1S AV-INCM~near  
 ‘The people, when will they approach me?’

In (6), we see that this topic position is not licensed by a complement clause. We can assume that the position of external topics is simply not present in the truncated structure which serves as a complement clause (Wurmbrand 2003). Note that topics are not generally banned from appearing in subordinate clauses but rather must appear in the internal topic position, following the complementizer and any interrogative elements, if present. The sentence in (7) demonstrates the simultaneous occupation of all the left-peripheral projections in (1) except for NegP.

- (6) \*alam=ko aŋ=maŋa=táo ay kuŋ kailan...  
 know=1S.GEN NOM=PL=person TOP COMP when
- (7) alam=ko ForceP[kuŋ IntP[kailan TopP[[aŋ=maŋa=táo] ay FocP[sa=ákin  
 know=1S.GEN COMP when NOM=PL=person TOP OBL=1s  
 TP[Ø-lá~lápít]]]]]  
 AV-INCM-near  
 ‘I know where the people will approach ME.’

The edge of the pronominal clitic domain lies on the border of IntP and ForceP. These clitics are unable to surface higher than IntP (modulo clitic climbing) but must surface after the first potential host within this domain. As shown by (8) and (9), delaying placement of clitics to TP when a potential host in the CP layer is present is ungrammatical.

- (8) Sa=Maníla[=sila] nag-áral[=\*sila] naŋ=áraw-áraw  
 OBL=M.=3P.NOM AV.BEG-study=3P.NOM GEN=day-day  
 ‘It’s in Manila that they studied every day.’

- (9) Saan/kailan/paano[=ka] p<um>unta[\*=ka]?  
 where/when/how=2S.NOM <AV.BEG>go  
 ‘Where/when/how did you go?’

Unlike oblique/adjunct questions, constituent questions, formed with *ano* ‘what’, *alin* ‘which’, *ilan* ‘how many’ or *sino* ‘who’, require a cleft-like structure with the interrogatives appearing in a predicate position rather than in CP, as discussed in detail by Potsdam (2006) for Malagasy. Under no circumstances can a clitic argument appear with the interrogative in such questions, as shown in (10). Rather, pronominal clitics are bounded by the case phrase which contains the predicate in such sentences.<sup>4</sup>

<sup>4</sup> The apparent clause boundedness of clitics in such constructions has occasionally been taken as evidence for their biclausal nature. In fact, clitic placement in such constructions is entirely parallel with that of the more canonical type of sentence with an aspectual predicate. In both types, the real generalization is that clitics are contained within their case phrase. A possessor clitic, like *ko* in (i), is also prevented from attaching to a host external to the case phrase of its origin.

- (i) T<in>ign-an[\*=ko] an=kotse[=ko]  
 <RL>look-LV=1S.GEN NOM=car  
 ‘My car was looked at.’ (Alternative only good for: ‘I looked at the car’)

The only Philippine languages that I have come across which allows argument clitics in cleft-like constructions to escape their domain are the Samar-Leyte and Camotes Island dialects of Eastern Bisayan as reported by Wolff (1967). Contrasting this with the situation in Cebuano, Wolff (1967:75) states:

“In the Camotes dialect and in SL Bisayan, the tendency to place short attribute elements [i.e. clitics –DK] at the beginning of the predicate is carried further: a short word which is a constituent of the subject [the *ang* phrase –DK] – that is, not the entire subject, but only an attribute of the subject – may optionally be placed following the first word of the predicate. In Cebuano (and in other Tagalog languages) such elements must be placed next to the rest of the subject.

Aku=man=situn an=p<in>alit *Samar-Leyte*  
 1S.NOM=EMPH=that NOM=<AV.PRF>buy  
 ‘I was the one who bought that.’

Aku=man=situn an=mi-palit *Camotes*  
 1S.NOM=EMPH=that NOM=AV.PRF-buy  
 ‘I was the one who bought that.’

Aku=man ang=mipalit niánà *Cebuano*  
 1S.NOM=EMPH NOM=AV.PRF-buy that  
 ‘I was the one who bought that.’

- (10) Ano[\*=mo] aŋ=g<in>awá-Ø[=mo]?  
 what=2S.GEN NOM=<BEG>do-PV  
 ‘What did you do?’

The interrogative *bákit* ‘why’ induces slightly exceptional syntax. Unlike the other interrogatives, pronominal clitics may optionally ignore the presence of *bakit* and attach to the following possible host, as exemplified in (11).<sup>5</sup>

- (11) Bákit[=ka] p<um>unta[=ka]?  
 why=2S.NOM <AV.BEG>go=2S.NOM  
 ‘Why did you go?’

---

<sup>5</sup> This exceptionality has a clear diachronic basis. Historically, *bakit* is derived from a conjoined construction involving the interrogative *bakin* and a conjunct complement headed by *at*. Originally, ‘why’ questions were simply formed with *bakin*. This is amply attested in literary texts dating from as late as the early 20<sup>th</sup> century. An example from a text published in 1899 is seen in (i). Note that the pronominal clitics follow the interrogative directly. (The original orthography is given in the first line.)

- (i) *¿Baquin ca na tañgay ng dagta ng lanca?*  
 Bákin=ka=na taŋay naŋ=dagtà naŋ=lanka?  
 why=2S.NOM=ALR carry GEN=sap GEN=lanka  
 ‘Why are you already carried away by the sap of the lanka fruit?’ (Francisco 1899)

Sometime in the late 19<sup>th</sup> century, the conjoined construction gained in popularity. This latter construction is well attested in texts from the same period and eventually overtook its simpler predecessor. It is exemplified in (ii), from a poetic text published in 1913.

- (ii) *...baquin at ipapatay reinang ualang casalanan*  
 ...bakin at i-pa~patay reyna=ŋ wala=ŋ kasalanan  
 why CONJ CV-INCM~kill queen=LNK N.EXT=LNK sin  
 ‘Why kill the guiltless queen?’ (Martinez 1913; Fansler & Unson 1916:256-7)

Because the conjunction in Tagalog can never host clitics and pronominal clitics can never move out of a conjunct, the clitics simply attached to the first available host within the second conjunct. In this older stage of the language the pattern found with ‘why’ questions was as in (iii) (a constructed example).

- (iii) Bákin at[\*=ka] p<um>unta[=ka]?  
 why CONJ=2S.NOM <AV.RL>go  
 ‘Why did you go?’

As the interrogative, *bakin*, and the conjunction, *at*, were reanalyzed as a single lexical item, *bakit*, ‘why’ questions fell in line with other interrogatives and *bakit* was able to host clitics.<sup>5</sup> The historically conservative pattern, however, still survives as an option, as seen in **Error! Reference source not found.**, and is ultimately the cause for the conflicting judgments in the literature (see Kroeger 1993:67 for an example).

Interrogative elements also function non-interrogatively in subordinate clauses, in which case pronominal clitics still attach to the interrogative. In (12), for instance, the subject clitic in the embedded clause must attach to the interrogative and cannot follow the predicate head.

- (12) Hindi=ko alam kuŋ saan[=siya] p<um>unta[\*=siya]  
 NEG=1S.GEN know COMP where=3S.NOM <AV.BEG>go  
 ‘I don’t know where he went.’

High adverbial clitics, like the question marker *ba*, can surface as high as the outer topic position, as shown in (13), where both clitics surface in their leftmost position.

- (13) Bukas=**ba** nanɣ=gabi ay Ø-sa~sayaw=**silá** nanɣ=pandango?  
 tomorrow=QM GEN=night TOP AV-INCM~dance=3P.NOM GEN=fandango  
 ‘Tomorrow night, will they dance a fandango?’ (S&O:429)

As discussed by S&O (p.429) and Anderson (2005:175), only a certain subset of clitics can appear within high topics. Focus adverbs such as *lámanɣ* ONLY, cannot. The picture in (1) in conjunction with the theory presented here offers a solution to some of the basic linear problems encountered with 2P elements. The feature bundles which are ultimately spelled out as 2P clitics are merged at IntP (for pronominals) and higher projections (for adverbials). They are then positioned according to how they best satisfy the constraint ranking. The high ranking \*WEAKSTART (PPh) will insure that clitics will not appear initial within their prosodic phrase.

Regarding why pronominal clitics attach to a CP layer rather than TP, we can follow Carstens (2003), who offers several pieces of evidence for understanding CP to be the actual locus of subject agreement features. The features of the genitive/ergative argument are associated with a lower layer, but raise to adjoin to the subject features

for the sake of maintaining the unity of the clitic cluster. In languages which do not value the unity of the clitic cluster, such as Agutaynen (see §4.6.2) we find that the genitive/ergative clitics stay closer to the predicate head while nominative clitics surface in the left periphery. Adverbials on the other hand simply attach to a phrase at which they can be interpreted semantically (Ernst 2002, Kaufman 2005).

### 5.3 Impenetrability and the Clitic Visibility Condition

In this section we will explore the impenetrability of focus fronted oblique and adjunct phrases, the main topic of the present chapter. The basic case of pronominal impenetrability can be seen in (14). The complex oblique phrase, which contains several possible clitic hosts becomes opaque to clitic placement when fronted to a focus position. The ungrammatical (14)b exemplifies clitic placement within the opaque constituent.

(14) a. {sa=ma-laki=ŋ      báhay na    iyon}=sila      naka-tira  
           OBL=ADJ-big=LNK house LNK that.NOM=3P.NOM STA-live  
           ‘They live in that large house.’

b. \*{sa=ma-laki=sila=ŋ      báhay na    iyon}      naka-tira  
           OBL=ADJ-big=3P.NOM=LNK house LNK that.NOM STA-live

As we saw earlier in §2.3.3, there are good grounds for rejecting a Halpernian solution to impenetrability in Tagalog. Recall that Halpern (1995) suggests post-word placement results from Prosodic Inversion when an enclitic has nothing to its left. According to Halpern, impenetrable phrases in the left periphery are simply above the domain of clitic positioning and thus clitics can stay in their base position without having to undergo inversion. This approach then clearly predicts that clitics should not move past an impenetrable phrase in the left periphery were one present. The

prediction is falsified by data such as in (15) where a clitic follows negation above a focus fronted oblique.<sup>6</sup>

- (15) Hindi=**ako** sa=ma-laki= $\eta$  syudad naka-tira  
 NEG=1S.NOM OBL=ADJ-big=LNK city STA-live  
 ‘I don’t live in a big bity’

We can thus dispense with a Halpernian solution to the impenetrability facts as its criterial prediction is not met. A purely morphological solution which stipulates the integrity of Focus Phrases, as in Anderson (2005), is seen to be inadequate as well as it fails to predict the fact that not all clitics are banned from appearing in focus fronted phrases; adverbial clitics which originate outside the focused phrase are perfectly acceptable in internal positions, a problem which is also common to purely prosodic analyses as that of Chung (2003). This can be seen in (16). Here, the 2P question marker *ba* can appear within the fronted focused oblique, shown in curly brackets, or after it. The pronominal clitic on the other hand can only appear following this phrase in the position indicated.

- (16) {sa=ma-laki[=**ba**]= $\eta$  báhay na iyon}[=**ba**]=kayo naka-tira?  
 OBL=ADJ-big=QM=LNK house LNK that.NOM=QM=2P.NOM STA-live  
 ‘Is it in that large house that you live?’

The dual positioning of clitics in (16), both after the first word and after the first syntactic phrase, instantiates a state of affairs explicitly ruled out by Halpern’s theory. Halpern treats “2P word” versus “2P phrase” as a parameterized choice for Prosodic Inversion and thus falsely predicts that their simultaneous satisfaction should

---

<sup>6</sup> It is worth noting that unlike Topic Phrases, the Focus Phrase is non-recursive, possibly related to the fact that it triggers a presuppositional reading of its complement (Kaufman 2005). It is thus impossible that sentences like (15) contain a “low focus” position.

be impossible. More generally, any approach which treats fronted obliques as de facto impermeable constituents will fail to explain the behavior of adverbial clitics in Tagalog.

We can rule out the *a priori* plausible analysis which would treat the question marker as sensitive to interrogative scope (cf. Japanese kakari particles as discussed by Whitman 1997). The question marker does not show much freedom of positioning, as shown by (17), a sentence with unmarked verb-initial order. Here, the question marker cannot appear within the oblique but must follow the clause-initial verb (the first legitimate host), regardless of interrogative scope. We are thus at odds to posit that the question marker originates within the oblique before movement.

- (17) naka-tira[=**ba**]=kayo sa=ma-laki[\*=**ba**]=ŋ báhay na iyon?  
 STA-live=QM=2P.NOM OBL=ADJ-big=QM=LNK house LNK that.NOM  
 ‘Do you live in that large house?’

It can also be imagined that interrogative scope must be indicated on an oblique by both fronting it to the focus position *and* association with the question marker. But this, too, is incorrect, as demonstrated by the fact that the question marker can appear with topics (S&O; Billings & Konopasky 2002; Anderson 2005), which are clearly not part of the interrogative scope at all, as seen in (18).

- (18) Bukas=**ba** naŋ=gabi ay Ø-sa~sayaw=sila naŋ=pandango?  
 tomorrow=QM GEN=night TOP AV-INCM~dance=3P.NOM GEN=fandango  
 ‘Tomorrow night, will they dance a fandango?’ (S&O)  
 (Not, ‘Will they dance a fandango TOMORROW NIGHT?’)

A successful approach to impenetrability, it is argued here, must crucially make reference to syntactic embeddedness while at the same time allowing 2P clitics to be positioned by surface criteria. Argument clitics, because of their status as dependents, must be visible to the predicate in the output. But visibility cannot be

defined in terms of standard c-command because, as shown in the evidence in chap.2, 2P clitics in Tagalog do not occupy a set syntactic position within the phrase structure. They are nonetheless *contained* by syntactic phrases and it is upon this fact which we can base our definition of visibility. The Clitic Visibility Condition upon which we will rely to derive the facts is defined in (19).

(19) Clitic Visibility Condition

For every argument clitic  $\alpha$  assigned a thematic role by a predicate head  $\beta$ , the minimal maximal projection *linearly containing*  $\alpha$  must dominate  $\beta$ .

This allows us to distinguish between different types of clitics based on their dependencies to other elements in the phrase structure in the familiar manner. Whereas argument clitics have a thematic relationship to an overt element in the syntax, adverbial clitics have no such relationship. Their licensing is thus more liberal than that of argument clitics in that they are only required to appear within the extended clause (as demarcated by IntP). Fronted obliques, although well within the placement domain of argument clitics, cannot embed such clitics precisely because they constitute independent syntactic phrases. Argument clitics are thus banned in such positions because they cannot “see” a predicate head outside the fronted oblique phrase.

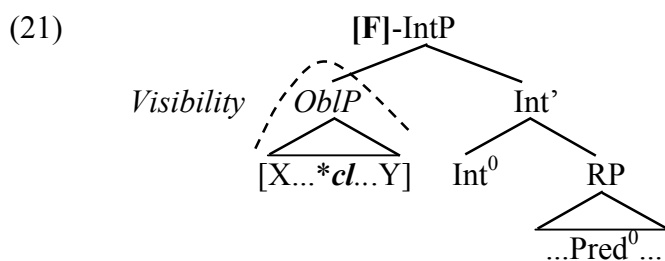
The notion of linear containment, italicized in (19) above, requires further elaboration. It, too, is a simple notion amounting to no more than unambiguous membership within a linear string. Linear containment is defined as in (20).

(20) Linear Containment

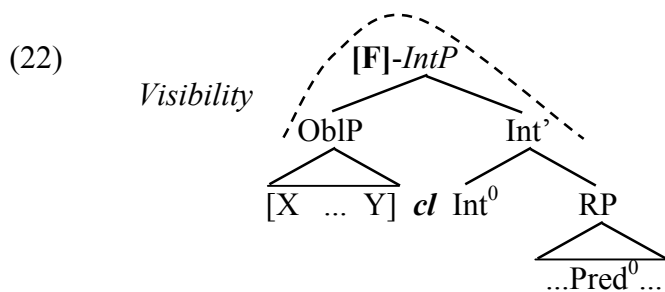
$\alpha$  is linearly contained in  $\beta$  if  $\alpha$  intervenes between overt terminal nodes dominated by  $\beta$



With the relevant definitions in place, we can now return to the banned structure in (21) where an argument clitic is contained within a fronted oblique. The clitic, sandwiched by syntactic terminals belonging to ObIP, is linearly contained by ObIP despite not occupying a syntactic position within it. By virtue of the Clitic Visibility Condition, ObIP marks the edges of the clitic's visibility domain. Because the verb is outside this domain, the condition is violated and the structure rendered illicit.

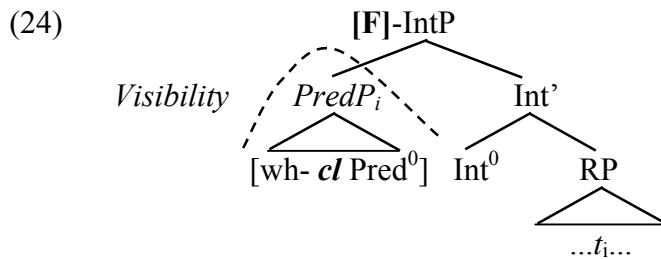


By contrast, clitics which are placed on the edge of ObIP are not linearly contained by it. Rather, the minimal maximal phrase which linearly contains clitics in this position in a structure like (22) is the IntP. Because IntP also dominates the predicate head, the visibility is maintained and the structure is grammatical. Furthermore, because internal topics, internal fronted obliques, auxiliaries, etc. are all housed in right-branching projections ultimately dominated by IntP as seen above in (1), no amount of intervening material in the left-periphery will be able to block the visibility between a clitic positioned after [Spec,IntP] as in (22) and its predicate.



This efficacy of the visibility condition is further demonstrated by the distinction found with fronted oblique phrases shown in (23). In (23)a, the oblique interrogative phrase contains the predicate by virtue of pied-piping. Therefore, even though the nominative clitic exceptionally appears embedded within [Spec,IntP], it is still visible because [Spec,IntP] also dominates the predicate head.<sup>7</sup> In contrast to this, when the predicate is external to the fronted oblique, as in the unmarked case, the clitic will not be able to appear embedded in [Spec,IntP]. This is what we find in (23)b, where the clitic must be visible to the predicate head *nagbigay* and thus cannot appear after the interrogative element. Oblique interrogative predicates have the basic structure shown in (24)

- (23) a. kaníno=**ka**=ŋ                    estudyánte?  
           OBL.who=2S.NOM=LNK student  
           ‘Whose student are you?’
- b. kaníno[\*=**ka**]=ŋ                    estudyánte[=**ka**] nag-bigay naŋ=pérà?  
           OBL.who=2S.NOM=LNK student=2S.NOM AV.BEG-give GEN=money  
           ‘To which student did you give money?’



<sup>7</sup> Note that (23) does not constitute a discontinuous wh- phrase, a structure which is categorically ungrammatical in Tagalog. This is clear from the fact that only one position exists for full-NP arguments in the same sentence, to wit, that of *Ernie* in (i).

- (i) Kanino=ŋ    estudyante si=Ernie?  
       OBL.who=LNK student P.NOM=E  
       ‘Whose student is Ernie?’

In the next subsection we see how the Clitic Visibility Condition applies to interactions between clitics and topics.

### 5.3.1 Topics

Tagalog permits topicalization of the nominative phrase, oblique phrases and certain adverbs to a left peripheral position followed by the topic marker *ay* (25), or comma-intonation (26). This has been seen in previous examples and has been discussed extensively in S&O, Kroeger (1998: chap.2) and Kaufman (2005) among others. Topics (and the topic marker) are strictly excluded from hosting pronominal clitics, as seen below.<sup>8</sup>

(25) Sa=Manílà[\*=sila] ay[\*=sila] nag-áral[=sila]            naṅ=áraw-áraw  
 OBL=M.                            TOP            AV.BEG-study=3P.NOM    GEN=day-day  
 ‘In Manila, they studied every day.’

(26) Sána[\*=ka], hindí[=ka]    ma-hiráp-an.  
 OPT                            NEG=2S.NOM    NVL-difficult-LV  
 ‘Hopefully, you will not find it difficult.’

In Kroeger’s (1993) analysis, topics are adjoined to IP. Topics which are set off by a pause instead of the topic marker are argued by Kroeger to be further out in the left periphery and analyzed as CP adjuncts. That clitics cannot attach to the right edge of left-peripheral topics has also been argued to fall out naturally from standard theories about the correspondence between syntactic and prosodic structure (Zec & Inkelas 1990, Truckenbrodt 1999 *inter alia*).<sup>9</sup> Because topics have to be phrased separately on the left edge of CP, they will always be strictly aligned to a prosodic

<sup>8</sup> Unlike Serbo-Croatian, the weight of the topic phrase is immaterial (Inkelas 1988, 1990; Radanović-Kocić 1988); a topic phrase in Tagalog can never be integrated into the clitic domain.

<sup>9</sup> Conravid & Kiparsky (2001), writing about Greek dialects, also note that, “For reasons which remain to be explored, cliticization is blocked across an adjunction boundary (topicalization and coordination - DK).”

phrase and clitics would then have to be phrased outside of their clause of origin. A problem arises here for such an approach to Tagalog as there also exist internal topics which appear within the normal bounds of clitic placement, as indicated by TopP in (1). Although it is very rare for internal topics to cooccur with interrogatives, we may observe an example in (27) (from a hymn in a classicizing register).<sup>10</sup>

- (27) *hangang kailan=kami sa=iyo ay mag-hi~hintay?*  
 until when=1+3.NOM OBL=2S TOP AV-INCM-wait  
 ‘Until when, for you, shall we wait?’<sup>11</sup>

The argument that topics are categorically outside of the clitic domain is thus not tenable as examples such as (27) demonstrate that clitics may appear to the left of internal topics given an appropriate host.<sup>12</sup> We must assume that internal topics can also appear without a preceding host such as the interrogative phrase in (27) and if this is the case then we need an independent basis for ruling out attachment to topics. That clitics cannot attach directly to [Spec,TopP] (e.g., as in the leftmost instance of *sila* in (25)) is predicted from Clitic Visibility Condition, but as we will see, this can only be part of the story. A clitic in the first position shown in (25) will be linearly contained within TopP and thus invisible to the projection line of the predicate. A similar explanation, however, is not available for why clitics cannot follow the topic marker. If the topic marker were proclitic like other functional heads, then the inability of other clitics to attach it may be explained by the fact that it is a weak monosyllabic element which is not prosodically deficient. However, the topic marker (along with the conjunction *at*) is exceptionally enclitic.<sup>13</sup> Because clitic stacking is clearly allowed in

<sup>10</sup> Non-attachment to the prepositional element *hangang* will be discussed in §4.2.

<sup>11</sup> From: <http://www.bukaspalad.com/board/showthread.php?p=16840>

<sup>12</sup> Internal topics differ from external topics in requiring the topic marker *ay*. External topics are only optionally marked with *ay* and can also be set off by intonational means.

<sup>13</sup> That the topic marker is enclitic is clear from its unique interactions with the segmental phonology of the word it follows. Following a vowel-final word, *ay* is reduced to *y* and following a word ending in *n* and occasionally *w*, the coda deletes. This is seen in (i), where the topicalized DP ending in *panahon*

Tagalog, we cannot then base our explanation on the prosodic make-up of the topic marker itself. Alternatively, it may be the case that topics are somehow special and are best analyzed as syntactic “orphans” (Shaer 2003, Espinal 1991), that is, “elements that are independent of their host sentences in the narrow syntax, and that, accordingly, have no hierarchical relation to them” (Shaer 2003:459). An orphan analysis of Tagalog topics may help to explain some of their other, largely unnoted, syntactic properties. For instance, unlike *wh*- operators or clefted material, topics can (i) be resumed by pronouns, as in (28),<sup>14</sup> (ii) escape strong islands, as in (29) (conditional island) and (30) (negative plus *wh*- island), (iii) seemingly allow extraction of the internal argument of actor voice verbs, as in (31), (iv) seemingly allow extraction out of the external argument of non-actor voice verbs, as in (32) and (33), and (v) disobey superiority, as shown by the minimal pair in (34), in which nominative and oblique arguments can both be topicalized in either order.

(28) Si=Juan, Ø-á~alis(=siya)  
 P.NOM=J AV-INCM~leave=3S.NOM  
 ‘Juan, (he) will leave.’

(29) Ikaw ay hindi=ko alam kuṅ=mag-tá~tapos o hindi  
 2S.NOM TOP NEG=1S.GEN know COMP=AV-INCM~finish or NEG  
 ‘You, I don’t know if (you) will graduate or not.’

---

‘time’, is phonologically reduced as shown. We can thus rule out a proclitic-like phonological dependency on the following material.

(i) <sub>PPh</sub>[aŋ]=<sub>PWd</sub>[panaho=y]] t<um>á~takbo=ŋ pa-layò  
 NOM=time=TOP <AV.RL>INCM~run=LNK DIR-far  
 ‘Time is running away.’ (cf. *panahon* ‘time’) (From “Harapin” by Spongecola)

Interestingly, while there is no prosodic dependency between the topic marker and following material, there appears to be a morphosyntactic one. No manner of ellipsis, movement or conjunction may strand the topic marker or otherwise separate it from its following clausal complement. In this, *ay* differs from familiar topic markers (e.g., Japanese *wa*) which do not form such a tight link with the following clause. This must be related to the diachronic origin of the topic marker as a proposed determiner on the following material.

<sup>14</sup> This is more natural with topics set off by a pause rather than those introduced with the marker *ay*, although resumption is also attested with the latter type.

- (30) kahit pag-sakay naṅ=*bus* ay hindi=ko alam kuṅ=papano  
 even GRND-ride GEN=*bus* TOP NEG=1S.GEN know COMP=how  
 ‘Even riding the bus, I didn’t know how.’<sup>15</sup>
- (31) ...at aṅ=*old testament*=naman=pò ay sino aṅ=*nag-sulat*?  
 and NOM=O.T.=SWTCH=POL TOP who NOM=AV.BEG-write  
 ‘...and the Old Testament who was the one who wrote (it)?’<sup>16</sup>
- (32) Káhit si=JESUS Christ ay hindi=táyo káya=ṅ i-*save*  
 even P.NOM=J.C. TOP NEG=1P.NOM able=LNK CV-save  
 ‘Even Jesus Christ, won’t be able to save us
- pag-datiṅ naṅ=pag-hu~húkom  
 GRND-arrive GEN=GRND-TR~judge  
 when judgment comes.’<sup>17</sup>
- (33) aṅ=hukbo naṅ=Katipúnan ay ano aṅ=g<in>awà-Ø  
 NOM=army GEN=K TOP what NOM=<BEG>do-PV
- naṅ=áraw na yáon?  
 GEN=day LNK that.NOM  
 ‘The army of the *Katipunan*, what did (they) do on that day?’<sup>18</sup>
- (34) a. Di=ko káya=ṅ **sa=ákin** ay **ika-y** l<um>ayò  
 NEG=1S.GEN able=LNK OBL=1S TOP 2S.NOM-TOP <AV>far  
 ‘I cannot withstand for you to distance yourself from me’<sup>19</sup>
- b. Di=ko káya=ṅ **ika-y** **sa=ákin** ay l<um>ayò  
 NEG=1S.GEN able=LNK 2S.NOM-TOP OBL=1S TOP <AV>far  
 ‘I cannot withstand for you to distance yourself from me’

Considering this large amount of freedom afforded to topics in comparison to other, clearer cases of movement in Tagalog (e.g., focused obliques), it is not unreasonable to analyze topics as orphans. It then follows that argument clitics which are parsed with orphan constituents must be interpreted with those constituents as

<sup>15</sup> From:

<http://www.jappinoy.com/index.php?s=e2d955edff88bb919c2ebc1cfd5f3b2&showtopic=1597>

<sup>16</sup> From: <http://www.naga.gov.ph/forum/index.php?topic=1346.50:wap2>

<sup>17</sup> From: <http://www.filipino.ca/forum/thread-view.asp?threadid=28970&start=1&posts=11>

<sup>18</sup> From: <http://www.tagalog.ws/mod/wiki/view.php?id=19&page=Ang+Paghihimagsik>

<sup>19</sup> From: <http://www.emanilapoetry.com/writersgroup/index.php/2007/11/20/para-sa-iyo/>



explain a class of apparent “clitic-third” sentences. The relevant optionality (without an overt topic marker) is shown in (36) and (37).

(36) (i-)p<in>a-úlet sa=ákin kuṅ=saan=*ako* hindì ma-galing.  
 (CV-)<BEG>CAU-repeat OBL=1S COMP=where=1S.NOM NEG ADJ-great  
 ‘I was made to repeat where I wasn’t great.’<sup>22</sup>

(37) ...nása=lugar=*ako* kuṅ=saan hindí=*ako* maka-galaw  
 OBL=place=1S.NOM COMP=where NEG=1S.NOM AV.ABL-move  
 ‘I’m in a place where I can’t move.’<sup>23</sup>

In the former sentence, the clitic appears where expected, following the interrogative. In the latter sentence, however, the clitic appears after negation despite the presence of the interrogative in what appears to be the same clitic domain. However, if sentences such as (37) represent structures in which non-interrogative *saan* has been topicalized, as in (35), these cases are brought back into line with our expectations given that topics are ineligible hosts for clitics.

### 5.3.3 Parentheticals

Parentheticals are of some importance in the evaluation of competing theories of second position. Bošković (2001) claims that 2P clitics in Serbo-Croatian are bounded by their intonational phrase and consequently cannot cross parentheticals. In Tagalog, oblique phrases such as *sa tinin ko* ‘in my view’ and *sa palagay ko* ‘in my opinion’, among others, are commonly employed as parentheticals, in a similar manner to the corresponding English translations. Unlike Bošković’s interpretation of the Serbo-Croatian data, pronominal clitics in Tagalog are clearly allowed to have trans-parenthetical dependencies when attaching to hosts further to the left. As shown

<sup>22</sup> From: [http://pondahan.pansitan.net/2006\\_10\\_01\\_kiwipinay\\_archive.html](http://pondahan.pansitan.net/2006_10_01_kiwipinay_archive.html)

<sup>23</sup> From: <http://cubicledeespeyal.blogspot.com/>



in (38)-(40), when a parenthetical separates a preceding potential host from the predicate; attachment to the predicate or to the parenthetical itself is strictly ungrammatical.<sup>24</sup>

(38) Hindi[=táyo], sa=palagay=ko[\*=táyo], maka~ká-ratiŋ[\*=táyo]  
 NEG=1+2P.NOM OBL=opinion=1S.GEN AV.ABL~INCM-arrive  
 ‘We won’t, in my opinion, be able to arrive.’

(39) Paano=**nyo**, sa=tiŋin=nyo, ma-kú~kumbinsi  
 how=3P.GEN OBL=sight=2P.GEN PV.ABL-INCM~convince  
 ‘How can you convince, in your opinion, the millions of

aŋ=milyon-milyoŋ Pilipino na i-bóto=siya?  
 NOM=million-million:LNK Filipinos COMP CV-vote=3S.NOM  
 Filipinos to vote for him?’<sup>25</sup>

(40) Saan=**ka**, sa=tiŋin=mo, Ø-pú~punta, Láŋit?  
 where=2S.NOM OBL=sight=2S.GEN AV-INCM~go heaven  
 ‘Where, in your opinion, will you go, Heaven?’<sup>26</sup>

Thus, although prosody plays an undeniable role in clitic placement, Bošković’s (2001) suggestion that clitics are restricted to appearing within their intonation phrase of origin cannot be correct for Tagalog.<sup>27</sup> In line with the present analysis, prosody disallows clitics to appear in phrase initial position but cannot create islands for clitic movement independent of syntax. The demarcation of clitic placement domains is defined in strictly syntactic terms.

<sup>24</sup> Although not exemplified here, clause initial parentheticals cannot host clitics and are thus similar to topics in this respect.

<sup>25</sup> From:

<http://www.chizescudero.com/phpBB2/viewtopic.php?p=1121&sid=f08b92f3c6aee2374d6323ceb13abf12>

<sup>26</sup> From: <http://www.yehey.com/boards/ShowPost.aspx?PostID=41119>

<sup>27</sup> Based on some casual informant work with native speakers, it seems that the Serbo-Croatian facts upon which this aspect of his proposal were based are not entirely clear either. It thus still remains to be seen if pure prosodic phrasing can “trap” clitics in lower positions.

#### 5.4 Clitic climbing

As stated above, the Clitic Visibility Condition refers to clitics and predicate heads which are in a thematic role assigning relationship. However, due to the familiar process of clitic climbing, clitics may also appear in clauses above those in which they are assigned a thematic role. Tagalog displays optional clitic climbing with the modals *dápat* ‘must’, *kailáhan* ‘need’, *puwéde* ‘can’, *káya* ‘able’ and *maaári* ‘can’.<sup>28</sup> As shown in (41), a nominative clitic associated with an embedded predicate may follow a matrix modal although this is impossible with non-clitic nominative arguments, as seen from (42).

- (41) **Gusto=ko[=siya]=ŋ**                      <um>alis[=siya]  
 want=1S.GEN=3S.NOM=LNK    <AV>arrive  
 ‘I want him to leave’
- (42) **Gusto=ko**    [\*aŋ=paŋúlo=ŋ]                      <um>alis    [aŋ=paŋúlo]  
 want=1S.GEN    NOM=president=LNK    <AV>leave    NOM=president  
 ‘I want the president to leave’

As can be expected, if negation or other pre-predicate elements precede the modal the clitic appears before the first legitimate host, as in (43).

- (43) **Hindi=ko=siya**                      gusto=ŋ    magiŋ    paŋúlo  
 NEG=1S.GEN=3S.NOM    want=LNK    AV:become    president  
 ‘I don’t want him to become the president.’

There has been considerable work on the relationship between restructuring and clitic climbing, beginning with Rizzi (1982). In several Romance and Slavic languages, clitic climbing has been shown to crucially depend on reduced structure in the lower clause as a result of restructuring. Kroeger (1993:181-5) makes the same

---

<sup>28</sup> In other language families for which clitic climbing has been investigated, climbing also typically occurs with causative verbs but causation is morphological and not syntactic in Philippine languages, and thus not a plausible candidate.

observation for Tagalog. A regular biclausal control structure with the matrix predicate *káya* able is shown in (44)a. *Káya* assigns genitive case to its agent and selects a complement clause, introduced by the linker/complementizer. The restructured version is shown in (44)b, where *káya* appears to form a complex predicate with the subordinate verb. As a result, the agent follows the subordinate verb and is assigned nominative case, as expected with an actor voice verb such as *bumili*.

- (44) a. *Káya ni=Manuel na b<um>ili naη=bágo=η kótse* (Control)  
 able P.GEN=M. LNK <AV>buy GEN=new=LNK car  
 ‘Manuel can buy a new car.’ (Kroeger 1993)
- b. *Káya=η b<um>ili si=Manuel naη=bágo=η kótse* (Restructuring)  
 able=LNK <AV>buy P.NOM=M. GEN=new=LNK car  
 ‘Manuel can buy a new car.’ (Kroeger 1993)

Because clitic climbing can only be observed independently with a transitive subordinate verb, Kroeger illustrates the facts with *utúsan* ‘to order’ in the locative voice. Here, the assignment of case will not differ between the upstairs and downstairs predicate but restructuring can still be diagnosed by the position of the agent. An agent following the matrix predicate and followed by the linker/complementizer indicates a fully biclausal structure while an agent following the lower predicate indicates restructuring has taken place. The examples in (45)a and b show restructuring with and without clitic climbing. As in Romance, restructuring necessitates clitic climbing, as seen by the ungrammaticality of (45)b. Without restructuring, the agent remains in the matrix clause and the clitic argument of the lower verb must remain in the subordinate clause, as in (45)c.

- (45) a. *Hindi=siya káya=η utús-an ni=Pedro* (Restructuring)  
 NEG=3S.NOM able=LNK order-LV P.GEN=P.  
 ‘Pedro cannot order her around.’ (Kroeger 1993:183)
- b. \**Hindi káya=η utús-an=siya ni=Pedro* (Restructuring)  
 NEG able=LNK order-LV=3S.NOM P.GEN=P. (Kroeger 1993:183)

- c. Hindi *káya ni=Pedro=η utús-an=siya* (Control)  
 NEG able P.GEN=P.=LNK order-LV=3S.NOM  
 ‘Pedro cannot order her around.’ (Kroeger 1993:183)

Kroeger concludes that there is no IP boundary in restructured clauses and that the clitic placement facts fall out naturally from this fact. In both cases, the clitic is positioned after the first  $X^0$  within the smallest IP. For the control and restructured clauses in (46), Kroeger (1993:189) posits the underlying structures in (47), with arrows indicating clitic movement to 2P from argument position.

- (46) a. Hindi *káya ni=Pedro bigy-an=siya naη=pérà* (Control)  
 NEG able P.GEN=P. give-LV=3S.NOM GEN=money  
 ‘Pedro is not able to give him money.’
- b. Hindi=*siya* *káya=η bigy-an ni=Pedro naη=pérà* (Restructuring)  
 NEG=3S.NOM able=LNK give-LV P.GEN=P. GEN=money  
 ‘Pedro is not able to give him money.’

- (47) a. IP[Hindi<sub>Γ</sub>[*káya* s[*Pedro* CP[C’[=η IP[s[*bigyan pera siya*]  
 NEG able Pedro LNK give-LV money 3S.NOM
- b. IP[Hindi<sub>Γ</sub>[*káya* s[=η s[*bigyan Pedro pera siya*]  
 NEG able LNK give-LV Pedro money 3S.NOM

The view of clitic climbing resulting from the lack of intervening structure is common to most analyses (cf. Wurmbrand 2001). It appears, however, that the reduction which takes place in Tagalog restructuring contexts cannot be all that far-reaching. Mercado (2002) observes several phenomena which are predicted to be absent under major reduction but are nonetheless permitted in Tagalog restructured clauses.<sup>29</sup> We additionally note here the possibility of oblique focus fronting in a

<sup>29</sup> It’s not clear to me that all of Mercado’s (2002) grammaticality judgments are universally accepted. In particular, Mercado claims that sentences such as (i) are acceptable, in which the higher predicate assigns genitive case to an argument in a lower intransitive clause, and (ii), in which clitic climbing

subordinate restructured clause. This should also be impossible if restructuring entails the formation of a complex predicate as no left-peripheral projections should be present in such a small domain. In (48), we see that a fronted oblique is able to intervene between the two predicates of a restructured clause and in (49) we see the same intervention with an adjective-headed verb phrase (argued by Kroeger 1993 to instantiate the same basic phenomenon as restructuring).

(48) *hindì káya=ŋ sa=ákin l<um>ában aŋ=paŋúlo*  
 NEG able=LNK OBL=1S <AV>fight NOM=president  
 ‘The president can’t fight AGAINST ME.’

(49) *Ma-dalas=siya sa=ákin mag-pa-print naŋ=maŋa=akdà=niya*  
 ADJ-often=3S.NOM OBL=1S:LNK AV-CAU-print GEN=PL=work=3S.GEN  
 ‘He often had ME print his work.’<sup>30</sup>

If focus fronting of obliques is possible with restructuring and clitic climbing then an analysis in which clitics seek the highest IP must at least be amended so that they seek a higher projection. The articulated left periphery approach may avoid this problem if we identify the phrase to which argument features adjoin as the interrogative phrase (IntP). If IntP is deleted in restructuring, the relevant features will be adjoined to the matrix IntP and clitics will thus be positioned in relation to the

---

takes place without restructuring. It is possible that the use of personal names obscures the distinction between clitics and full NPs in as personal names behave as optional clitics in Tagalog (Billings 2005). If this is the case, (ii) simply represents clitic climbing of both arguments.

- (i) (?)*Gusto=ŋ s<um>ayaw ni=Isabel si=Pedro*  
 want=LNK <AV>dance P.GEN=I. P.NOM=P.  
 ‘Isabel wants Pedro to dance.’
- (ii) (?)*Gusto=siya ni=Isabel na s<um>ayaw*  
 want=3S.NOM P.GEN=I. LNK <AV>dance  
 ‘Isabel wants him to dance.’

<sup>30</sup> From: <http://richardernaciopage.blogspot.com/2007/12/mag-alaala-ko-kay-sir-rene-sa-up-icw-at.html>

higher clause. That no IntP exists in the lower clause is independently verifiable from the ungrammaticality of (50) and (51) below, in which the focused obliques in (49) and (48) have been replaced by an oblique interrogative.

(50) \*hindi káya=ŋ **kanino** l<um>ában aŋ=paŋúlo  
 NEG able=LNK OBL:who <AV>fight NOM=president

(51) \*Ma-dalas=siya **kanino** mag-pa-*print* naŋ=maŋa=akdà=niya?  
 ADJ-often=3S.NOM OBL:who AV-CAU-print GEN=PL=work=3S.GEN

By not projecting IntP, both clauses are contained within a single phase. Clitic movement to the higher predicate thus still respects the Phase Correspondence Principle. The Clitic Visibility Condition will also be satisfied by clitic climbing so long as the clitics in the higher clause are linearly contained by a phrase on the same projection line as the lower predicate.<sup>31</sup> We are thus able to predict without stipulation that clitics can appear in a higher clause when belonging to a truncated embedded clause. Crucially, this is predicted without reference to syntactic movement thereby maintaining the generalization that full NP arguments cannot raise in the same way that adjoined clitics can.

---

<sup>31</sup> As expected from the feature adjunction approach taken here, the dependency between raised clitics and verb in the lower clause cannot be blocked by overt subjects in the subordinate clause as suggested for French clitics by Kayne (1975). Clitic climbing is licensed in (i) despite the presence of an overt external argument in the embedded clause:

(i) Gusto=ko=**siya**=ŋ ma-kítà naŋ=maŋa=*officemates*=ko.  
 want=1S.GEN=3S.NOM=LNK PV.NVL-see GEN=PL=officemate=1S.GEN  
 'I want my officemates to see her.'<sup>31</sup>

Incidentally, this also suggests that it is the finiteness of the verb which is responsible for blocking effects rather than the presence of a subject, as suggested in several works attempting to reduce clause-boundedness to Relativized Minimality, or related locality conditions. Notably, Tagalog differs from other languages typically observed for clitic climbing in not showing any dependencies between tense/aspect and case assignment.

## 5.5 Clitic positioning within DP and non-verbal predicates

In this section we will demonstrate that clitic positioning follows the same principles across domains, including DP internally. Among those environments examined here are nominal phrases, adjectival phrases and prepositional phrases. A major difference between non-aspectual and aspectual predicates is that the former is more sensitive to the role of referentiality. As will be shown in the following sections, referentiality effects clitic placement in an interesting way in Tagalog.

### 5.5.1 *Nominal constituents*

In this section we examine three aspects of clitic placement in nominal constituents: (i) the placement of clitics within X'-constituents, namely, nominal constituents containing a head plus modifiers, (ii) the outer bounds of clitic placement in the nominal domain, (iii) the role of referentiality in clitic placement.

#### 5.5.1.1 *Modified nominal constituents*

In complex NPs with phrasal possessors, the clitic must attach directly to the nominal head in the now familiar pattern. Attachment to the right edge of the larger constituent is ungrammatical, as seen in (52).

- (52) Anak[=ka]    nan=pulis[\*=ka]  
child=2S.NOM GEN=police=2S.NOM  
'You're the child of a police officer.'

However, as noted by Kroeger (1993), in the structures he analyzes as X' constituents, pronominal clitics only optionally intrude; positioning after an N+Adj constituent is also possible, as shown in (53). Again, the first position is unavailable for full NPs (54).

- (53) Ma-laki[=ka]=ŋ táo[=ka]  
 ADJ-big=2S.NOM=LNK person=2S.NOM  
 ‘You’re a big person.’
- (54) Ma-laki [\*aŋ=paŋúlo]=ŋ táo [aŋ=paŋúlo]  
 ADJ-big NOM=president=LNK person NOM=president  
 ‘The president is a big person.’

We return to these facts in §5.4.1.3 in the discussion of referentiality. For now, we concentrate on the fact that the freedom of positioning in (53) does not apply to all combinations of nouns and modifiers. Tagalog allows permutation of certain elements within the DP. This can be seen in the alternative orderings for a simple ADJ+NOUN phrase in (55)a-b (see Kaufman 2006 for some conditions on NP internal scrambling). However, there exists an asymmetry in the positioning possibilities of clitics between the adjective and nominal initial orders, as can be seen by the comparison of (56) with (53). When the adjective is postposed, as in (56), argument clitics cannot follow it (Donohue n.d., Bernd 2002).

- (55) a. ma-laki=ŋ táo                      b. táo=ŋ              ma-laki  
 ADJ-big=LNK person                      person=LNK ADJ-big  
 ‘big person’                                      ‘big person’
- (56) táo[=siya]=ŋ              ma-laki[\*=siya]  
 person=3S.NOM=LNK ADJ-big=3S.NOM  
 ‘She’s a big person.’

As Richards (1999) notes, there also exists an asymmetry in the noun initial and modifier initial orders such that only the former may optionally employ the post-consonantal/pausal linker *na* even if the preceding segment is a vowel. To demonstrate with the previous example, both orders allow the two elements within the





The facts could offer additional support to the analysis posited in Kaufman (2005) which views such reorderings as the output of prosodically driven movement (Zubizarreta 1998) aiming to align a relatively fixed (right-aligned) intonational peak with a focused element (cf. Huck & Na 1990 for a similar relation between focus and extraposition in English). Because, only one intonational peak is licensed within the DP domain under normal circumstances, it follows that the intonational alignment process can only motivate a single rightwards scrambling. Additional scrambling, as in (58)c and d, could not serve the purpose of aligning focused modifiers with the natural locus of intonational prominence.

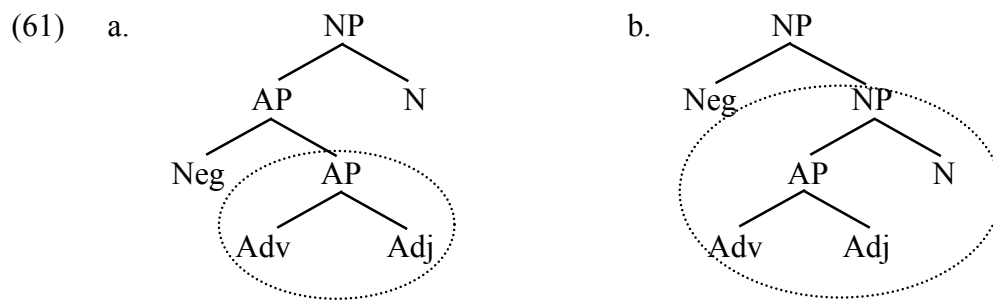
The Clitic Visibility Condition is also responsible for interesting asymmetries between nominative and genitive clitics in complex NPs with regard to certain intermediate positions in the DP. While the intermediate clitic positions shown in (59) are ungrammatical for nominative clitics in predications, some are only moderately degraded for possessor clitics within the DP, as seen in (60).

(59) hindi[=ako] masyado[\*=ako]=η ma-talino[\*=ako]=η estudyante[\*=ako]  
 NEG=1S.NOM overly =LNK ADJ-smart =LNK student  
 ‘I’m not too smart a student.’

(60) siya aη=hindi[=ko] masyado[\*=ko]=η ma-talino[<sup>?</sup>=ko]=η  
 3S.NOM NOM=NEG=1S.GEN overly =LNK ADJ-smart =LNK  
 estudyante[<sup>?</sup>=ko]  
 student  
 ‘He’s my not too smart student’

This is a direct result of differing syntactic configurations and the possibilities they leave for clitics to comply with the Clitic Visibility Condition. In particular, elements like negation can attach at various points within nominal structures, negating

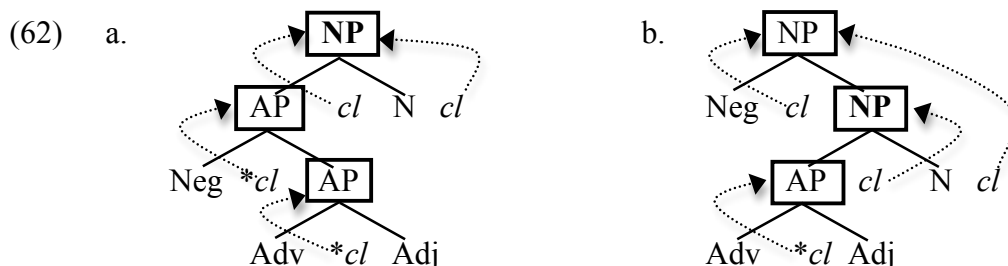
smaller or larger constituents.<sup>32</sup> Sentential negation on the other hand, is more restricted in its appearance, as it can only take proper scope by attaching to a high clausal projection. Let us consider first the two representations in (61) of the NP of interest in (60). I assume here that negation takes scope via association with focus within its c-command domain (Rooth 1985). For negation to scope over the adjective (as it must for many speakers in this sentence), negation can either attach low, as in (61)a, or high as in (61)b. Both are possible so long as the scope of negation is included within the c-command domain, indicated by the dotted circle.



Now consider what happens when we add possessor clitics to these two different structures, as seen in (62), where *cl* indicates all the possible clitic positions within the string and the arrows indicate the phrase within which they are linearly contained from a given position. According to the clitic visibility condition, the minimal maximal projection which linearly contains clitics must also dominate the predicate head with which they are associated. Possessor clitics associate with their head noun and thus cannot be trapped within complex adjuncts to the left of the noun. In (62)a, when negation adjoins low, clitics cannot follow negation or any AP internal material and still be visible to the predicate head, thus incurring the stars seen on the two first clitic positions. On the other hand, when negation attaches high, as in (62)b, a

<sup>32</sup> We treat both nominal negation and adjectival modification as adjunction here for convenience although the results may not differ significantly if these are treated as functional projections.

possessor clitic may follow negation and still be visible to the predicate head below. AP internal positions are again ruled out, as in the previous structure.



When we further take into consideration the clitics' need to appear as close to the site of feature adjunction as possible, we are left with two possibilities for clitic positioning here, correlating to the two sites of negative adjunction. With low adjunction, the clitic will follow the entire AP and with high adjunction it will follow negation itself. Low adjunction can thus force apparent delayed clitic placement. Because low adjunction is only a possibility with non-sentential negation, this gives the appearance that genitive clitics within DP are freer than argument clitics in the clausal domain. This variation should thus not be mistaken for simple freedom but rather is a consequence of two different syntactic structures. Note, however that there exists another position available for possessor clitics within the DP corresponding to the final bracketed clitic in (60). This possibility represents a further asymmetry between DP and clausal clitics and is to be explained by independent semantic factors, discussed in §5.4.1.3.

### 5.5.1.2 *The left edge of the nominal clitic domain*

Identifying a unique nominal/DP domain in Tagalog is not a trivial task as the distinction between nominal and verbal syntax is extremely subtle (Himmelman

1997, 2008, Kaufman 2009). Elements which are typically considered as belonging to both the verbal and nominal domains can to a large extent comingle within complex constructions. A full discussion of the problems this involves would take us too far afield and thus we continue to employ the term nominal domain to refer to argument constituents and to predicates headed by nouns without implying that such a domain strictly excludes typically verbal/clausal projections.

Consider the two structures in (63)a and b, which only differ on the surface in that the nominal in (63)a is not case marked while that in (63)b is. In the former, it is possible to position a genitive (possessor) clitic directly after negation while in (63)b the clitic is “trapped” within the case phrase.

- (63) a. hindi=ko anak  
 NEG=1S.GEN child  
 ‘not a child of mine’
- b. hindi[\*=ko] arj=anak[=ko]  
 NEG=1S.GEN NOM=child  
 ‘not my child’

Recall from our discussion of  $K^w ak^w ala$  that the DP appeared to constitute a phase for the purpose of clitic placement. Clitics adjoined to NP could not encliticize to an element in the D layer, outside of the Spell-Out domain. That a similar principle is at work in Tagalog is clear from the behavior of demonstrative as selective clitic hosts. The actual edge of the relevant domain is the Case Phrase (KP following Fukui 1986, Lamontagne & Travis 1986), but because case markers are monosyllabic proclitics themselves, it is impossible to verify their ability to host other clitics, as they are ruled out on independent prosodic grounds. For this reason, demonstratives will be used to illustrate. Demonstratives are generated in the immediately lower projection

but raise to Kase<sup>0</sup>, as will be argued below, and thus sit outside the relevant Spell-Out domain.

Table 5.1, shows the demonstratives in Tagalog arranged by case and proximity features. Demonstratives are not only prosodic words but demonstrably prosodic word heads, as they can be stressed and may undergo intensive reduplication (e.g., *iyun~iyon* that:LNK~that ‘really that’).

**Table 5.1.** Tagalog demonstratives

	NOM/DEFAULT	GEN	OBL/DEICTIC
Speaker proximate	<i>ito</i>	<i>nito</i>	<i>dito</i>
Addressee proximate	<i>iyán</i>	<i>niyan</i>	<i>diyan</i>
Distal	<i>iyon</i>	<i>noon</i>	<i>doon</i>

As shown in (64), demonstratives can “frame” the DP they associate with, appearing both to its left and to its right, a phenomenon which is also prominent in Malagasy, among other Austronesian languages.

- (64) Ito=ŋ            áso=ŋ    ito  
 this.NOM=LNK dog=LNK this  
 ‘this dog’

Demonstratives are case marked for one of the three cases, as seen above in Table 5.1. The nominative column is also marked DEFAULT because it can be linked to non-nominative NPs when postposed. For instance, in (65)a, the NOM/DEFAULT demonstrative *ito* modifies a genitive object in its postposed position and in (65)b, an oblique object. Case agreement with the NP is ungrammatical in postposed position as shown by the impossibility of the genitive marked *nito* and oblique marked *dito* in these positions. Rather, as shown in (65)c, case marked demonstratives can only be used in preposed position.

- (65) a. Sino aŋ=k<um>úha      naŋ=kániŋ      (\*n-)ito?  
 who NOM=<AV.BEG>take GEN=rice:LNK (GEN-)this  
 ‘Who took this rice?’
- b. Sino aŋ=k<um>úha      sa=kániŋ      (\*d-)ito?  
 who NOM=<AV.BEG>take GEN=rice:LNK (OBL-)this  
 ‘Who took from this rice?’
- c. Sino aŋ=k<um>úha      nito=ŋ      kániŋ      (\*n-)ito?  
 who NOM=<AV.BEG>take GEN:this=LNK rice:LNK (\*GEN-)this  
 ‘Who took (from) this rice?’

When case marked demonstratives are employed, they replace the case markers. In fact, case marked demonstratives are the result of (obligatory) morphological merger between the case marker and the NOM/DEFAULT demonstrative, as in (66). Merger of demonstratives with the nominative case marker, shown in (66)c, results in deletion/zero exponence of the case marker.

- (66) a. *naŋ + ito* → *nito*                      b. *sa + ito* → *dito*  
 GEN this this.GEN                      OBL this this.OBL
- c. *aŋ + ito* → *ito*  
 NOM this this.NOM

As already suggested by the data in (65), it appears that there exists two positions for demonstratives in the Tagalog DP, a higher position from which the demonstrative merges with the case marker and a lower position where it does not. Both positions may be filled by the demonstrative given that they are not immediately adjacent to each other. For morpho-phonological reasons, the demonstrative must appear in the lower position when the plural marker is used; the plural marker is proclitic and cannot appear without a host to its right, as seen by (67)a. That the demonstrative to the right of the plural marker is indeed sitting in the lower position is

also clear from the fact that the case marker can merge with the high demonstrative above the plural marker, as shown in (67)b. Note also that, as shown in (67)c, merger between the case marker and the demonstrative cannot operate downwards with the lower demonstrative such that the fused morpheme appears after plural marking. Descriptively speaking, all arguments must be case initial, ruling out such an ordering on the basis of surface criteria.

- (67) a. Sino aŋ=k<um>áin      naŋ=maŋa\*(=ito)?  
 who NOM=<AV.BEG>eat GEN=PL=this  
 ‘Who ate these?’
- b. Sino aŋ=k<um>áin      nito=ŋ=maŋa=ito?  
 who NOM=<AV.BEG>eat this.GEN=LNK=PL=this  
 ‘Who ate these?’
- c. \*Sino aŋ=k<um>áin      maŋa=nito?  
 who NOM=<AV.BEG>eat PL=this.GEN

The lower demonstrative is in near absolute final position in the DP; it may only be followed by phrasal possessors. The demonstrative cannot appear between the plural marker and the nominal head as shown in (68). If following the plural marker, the demonstrative can only appear following the head noun, i.e., in final position.

- (68) Sino aŋ=k<um>áin      naŋ=maŋa[\*=ito=ŋ] tsokolate[=ŋ ito]?  
 who NOM=<AV.BEG>eat GEN=PL=this=LNK chocolate=LNK this  
 ‘Who ate these chocolates?’

What is important for our purposes here is that in neither position can the demonstrative host pronominal clitics, as shown in (69). There is no morphophonological basis for this as demonstratives show all signs of being prosodic word heads. The most compelling evidence that domains are at issue here comes from the fact that demonstratives are perfectly good hosts for adverbial clitics, as shown in



(70)-(72), where *ba* QM, *naman* SWTCH and *pala* MIR, respectively, are positioned directly after the demonstrative.

(69) ito[\*=ko]=η      dalawa[=ko]=η      anak[=ko]  
 this=1S.GEN=LNK two                      =LNK child  
 ‘These two children of mine.’

(70) Ito=**ba**=η      si=Geronimo na      s<in>á~sábi-Ø=mo...  
 this=QM=LNK P.NOM=G      LNK <BEG>INCM~say-PV=2S.GEN  
 ‘This Geronimo that you mention...’<sup>33</sup>

(71) Ito=**namaη**      naη-yá~yári      sa=*career*=ko      ηayon...  
 this=SWTCH:LNK AV.BEG-INCM~happen OBL=*career*=1S.GEN now  
 ‘This (thing) on the other hand happening to my career now...’<sup>34</sup>

(72) Ma-sakit isip-in na ito=**pala**=η      AFP ay...  
 ADJ-pain think-PV LNK this=MIR=LNK AFP TOP  
 ‘It’s painful to think that this AFP surprisingly is...’<sup>35</sup>

On analogy with crosslinguistic evidence and as suggested by the order of frozen morphemes within the fused case+ demonstrative morphemes (in particular, the *n*-formant on the genitive demonstratives), case precedes the demonstrative. Because of their complementary distribution we have posited that Kase<sup>0</sup> hosts the demonstrative head on the left periphery of the DP. The order of elements in (73)a summarizes the relevant DP architecture for present purposes. The inner bracketed range represents the domain of pronominal clitic placement and the outer brackets demarcate the range of adverbial clitics. The example in (73)b shows a DP instantiating most of the categories simultaneously.<sup>36</sup>

<sup>33</sup> From: [www.lawphil.net/judjuris/juri2004/jun2004/gr\\_150501\\_2004.html](http://www.lawphil.net/judjuris/juri2004/jun2004/gr_150501_2004.html)

<sup>34</sup> From: [www.malaya.com.ph/nov08/ente\\_2.htm](http://www.malaya.com.ph/nov08/ente_2.htm)

<sup>35</sup> From: [www.ellentordesillas.com/?p=2316](http://www.ellentordesillas.com/?p=2316)

<sup>36</sup> The pre-numeral position for plurality (not illustrated in (73)b) indicates an approximative morpheme while the post-numeral position represents plurality proper.

- (73) a. CASE -Adv [DEM-Pron [POSS-(PL)- NUM-PL-ADJ-N]]-DEM  
           ↑                  |  
 b. ito=ŋ      akiŋ              sampu=ŋ    maŋa=ma-laki=ŋ    áso=ŋ    ito  
     this=LNK   1S.GEN:LNK    ten=LNK    PL=ADJ-big=LNK    dog=LNK    this  
     ‘these ten big dogs of mine’

Just as in K<sup>w</sup>ak<sup>w</sup>ala, the D head is outside the Spell-Out domain of the DP phase and thus outside the domain of clisis. Possessor features are adjoined below KP, presumably at the same location where we find free genitive pronominal possessors in (73). Adverbial features, in contrast, can be adjoined higher and can thus appear above the Spell-Out domain directly attached to the demonstratives.

### 5.5.1.3 Referentiality and impenetrability

We saw above that pronominal clitics (both genitive possessors and nominative subjects) followed the first element within their domain. Looking within putative X’ domains, Kroeger (1993:153 fn.21) observes that there appears to be an asymmetry between clitic positioning with modified nominals and modified verbs in Tagalog: “My impression is that the order...in which the clitic appears following the X’ constituent, is reasonably natural in N’ (Adj + N) constructions but relatively disfavored in V’ (Adv + V) constructions”. Despite the presence of optionality, the possibility of apparent “delayed 2P” in the nominal case will be shown here to have a principled basis rooted in the semantics. Specifically, presupposed, or D-linked noun phrases permit (but do not require) delayed clitic placement.<sup>37</sup>

<sup>37</sup> Semantic distinctions are also reported by Alexiadou & Stavrou (2000) for Greek 2P genitive clitics within the DP. The Greek facts are different, however, as delayed position is claimed to allow multiple readings while the internal position, only one. In the Tagalog facts discussed below the internal position appears to be the default, with the external position allowing a subset of interpretations, although further work is necessary to clarify this point.

The minimal pair in (74) demonstrates the role of presuppositionality in clitic positioning. In the (a) continuation, the possessor clitic follows the first available host and the interpretation of the nominal is non-presuppositional. In the (b) continuation, the clitic is suspended after the nominal head and it is interpreted presuppositionally, i.e., the speaker has kids and does not like them.<sup>38</sup>

(74) Mahilig=ako sa=maña=bátà...  
 like=1S.NOM OBL=PL=kid  
 ‘I like kids...’

- |   |  |
|---|--|
| <p>a. basta hindi=<b>ko</b> anak<br/>         just NEG=1S.GEN child<br/>         ‘so long as they’re not kids of mine.’</p> | <p>b. basta hindi anak=<b>ko</b><br/>         just NEG child=1S.GEN<br/>         ‘just not my kids.’</p> |
|---|--|

The presuppositionality difference above is not restricted to nominals but can also be found with aspectual predicates. Observe the ambiguity in the sentence in (75), as represented by the differing English translations.

(75) Hindi s<in>úlat-∅ ni=Maligaya aη=gusto=nila=η ma-kítà  
 NEG <BEG>write-PV P.GEN=R. NOM=want=3P.GEN=LNK PV.NVL-see  
 ‘Maligaya didn’t write what they want to see.’ (non-presup. pred.)  
 ‘It’s not what Maligaya wrote that they want to see.’ (presup. pred.)

These two readings can be summarized as in (76)a and b. The existential force of the first term,  $\exists x[(\text{read})\text{speaker},x]$ , originates from the definite semantics of the nominative case marker *aη* while the existential force of the second term in (76)b correlates with the delayed clitic position, as seen earlier in (74)b.

(76) a. Non-presuppositional reading:  
 $\exists x[(\text{want.to.see})\text{they},x] \wedge \neg [(\text{wrote})\text{Maligaya},x]$

---

<sup>38</sup> Although the effects reported in this subsection are subtle and subject to variation between speakers, the difference between (74)a and b appears to elicit general agreement.

b. Presuppositional reading:

$$\exists x[(\text{want.to.see})_{\text{they},x}] \wedge \exists y[(\text{wrote})_{\text{Maligaya},y}] \wedge x \neq y$$

The two readings are disambiguated when we replace the full genitive agent in (75) with a clitic, as in (77). Now the sentence in (77)a corresponds unambiguously to the reading in (76)a, and that in (77)b, to that in (76)b.

(77) a. Hindi=**mo** s<in>úlat-Ø aŋ=gusto=nila=ŋ ma-kítà  
 NEG=2S.GEN <BEG>write-PV NOM=want=3P.GEN PV.NVL-see  
 ‘You didn’t write what they want to see.’ (non-presupp. pred.)

b. Hindi s<in>úlat-Ø=**mo** aŋ=gusto=nila=ŋ ma-kítà  
 NEG <BEG>write-PV=2S.GEN NOM=want=3P.GEN PV.NVL-see  
 ‘It’s not what you wrote that they want to see.’ (presupp. pred.)

This semantic distinction also explains a difference in the preference of clitic placement with two otherwise syntactically identical interrogative elements, *ilan* ‘how many’ and *alin* ‘which’. When the former is part of a complex interrogative NP in predicate position, pronominal clitics are preferably placed directly after the *ilan*, as shown in (78). When the same complex NP contains the D-linked interrogative *alin*, however, pronominal clitics are preferably positioned after the entire phrase, as shown in (79). That this difference in clitic placement is semantically, and not morphologically, grounded can be seen by the fact that *alin*, is a perfectly good clitic host for non-pronominal clitics. In (80), *alin* hosts the 2P incompletive and question marking clitics, *pa* and *ba*. It is thus clearly not the case that there is simply a tighter morphological bond between *alin* and the following material.

(78) ilan[=mo]=ŋ anak[?=mo] aŋ=naka-pások sa=UP?  
 how.many=2S.GEN=LNK child NOM=AV.NVL.BEG-enter OBL=UP  
 ‘How many of your children got into U.P.?’

(79) alin[?=mo]=ŋ anak[=mo] aŋ=naka-pások sa=UP?  
 which=2S.GEN=LNK child NOM=AV.NVL.BEG-enter OBL=UP  
 ‘Which child of yours got into U.P.?’

- (80) alin=**pa=ba**=η ahénsya nan=gobyérno aη=ma-tinò ηayon?  
 which=INCM=QM=LNK agency GEN=government NOM=ADJ-rational now  
 ‘Which other agency of the government is rational now?’<sup>39</sup>

This fact is also emergent in the very consistent tendencies for delayed placement with quantifiers versus cardinal modifiers.<sup>40</sup> The quantifier *báwat* ‘each’, while syntactically able to host clitics without drawing particularly unfavorable judgments from speakers, triggers delayed clitic placement when modifying an NP. This was tested by Google searches with the term *báwat kilos* ‘every move’ in addition to a possessive pronoun, i.e., ‘your every move’, ‘my every move’, etc. in both internal and post-nominal position. The results are shown in (81). With each of the four pronouns tested, the numbers show a highly significant preference for the post-nominal position.

- |         |                 |   |                          |    |                 |   |              |
|---------|-----------------|---|--------------------------|----|-----------------|---|--------------|
|         | 2 hits          | < | 109 hits                 |    | 0 hits          | < | 186 hits     |
| (81) a. | báwat[=ko=η]    |   | kílos[=ko] <sup>41</sup> | b. | báwat[=mo=η]    |   | kílos[=mo]   |
|         | each=1S.GEN=LNK |   | move                     |    | each=2S.GEN=LNK |   | move         |
|         | 4 hits          | < | 65 hits                  |    | 1 hit           | < | 43 hits      |
| c.      | báwat[=niya=η]  |   | kílos[=niya]             | d. | báwat[=nila=η]  |   | kílos[=nila] |
|         | each=3S.GEN=LNK |   | move                     |    | each=3P.GEN=LNK |   | move         |

This is in stark contrast to the situation found with cardinal modifiers, which are not necessarily presuppositional (Diesing 1992). Here, the distribution is more even but there is a consistent tendency for internal positioning of the possessor clitics rather than post-nominal positioning. The same set of pronouns were tested with the results shown in (82).

<sup>39</sup> From: <http://www.ellentordesillas.com/?p=132>

<sup>40</sup> Further research is required to address the important question of whether strong and weak quantifiers differ in this regard. The expectation is that weak quantifiers should behave like cardinal modifiers, preferring the internal position for clitics.

<sup>41</sup> Repeated hits of *bawat kilos ko* from transcriptions of a popular song were filtered out by subtracting the surrounding lyrics in the search.

- (82) a.  $\begin{array}{l} 307 \text{ hits} > 84 \text{ hits} \\ \text{dalawa[=ko]=}\eta & \text{kapatid[=ko]} \\ \text{two=1S.GEN=LNK} & \text{sibling} \end{array}$       b.  $\begin{array}{l} 23 \text{ hits} > 8 \text{ hits} \\ \text{dalawa[=mo]=}\eta & \text{kapatid[=mo]} \\ \text{two=2S.GEN=LNK} & \text{sibling} \end{array}$
- c.  $\begin{array}{l} 56 \text{ hits} > 19 \text{ hits} \\ \text{dalawa[=niya]=}\eta & \text{kapatid[=niya]} \\ \text{two=3S.GEN=LNK} & \text{sibling} \end{array}$       d.  $\begin{array}{l} 3 \text{ hits} > 1 \text{ hit} \\ \text{dalawa[=nila]=}\eta & \text{kapatid[=niya]} \\ \text{two=3P.GEN=LNK} & \text{sibling} \end{array}$

Why should presuppositionality effect clitic placement? It has long been noted that the possibility of subextraction from an argument is highly dependent on its presuppositionality or givenness. Erteschick-Shir (1973) refers to the semantic novelty required to license subextraction as “dominance”, offering the definition in (83)a and the constraint in (83)b. This is formulated in different terms by Diesing (1992) as the presuppositional NP constraint in (84) (see also, Fiengo and Higginbotham 1981).<sup>42</sup>

- (83) a. A clause or phrase is semantically dominant if it is not presupposed and does not have contextual reference (Erteschick-Shir 1973:22)
- b. Extraction can occur only out of clauses or phrases which can be considered dominant in some context. (Erteschick-Shir 1973:27)
- (84) *Presuppositional NP Constraint* (Diesing 1992:103)  
Extraction cannot take place out of a presuppositional NP

Diesing reinterprets the purely syntactic constraints argued by Horn (1974) and Bowers (1988) as semantically based, following a path laid out by Erteschick-Shir (1973). Diesing shows that the possibility of subextraction from NP correlates well with its possibility of obtaining an existential reading. NPs which cannot obtain an existential reading can also not be extracted from. For instance, objects of experiencer verbs cannot be interpreted existentially in the presence of quantificational adverbs

<sup>42</sup> Diesing ultimately revises this definition so that it follows from her tree-splitting hypothesis, but the formulation in (84) is sufficient for present purposes.

(Diesing 1992:chap.4) and this correlates well with the ungrammaticality of sub-extraction from such objects, as shown in (85).

(85) \*What do you usually like a picture of?

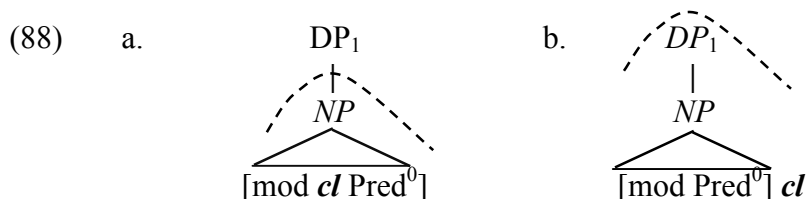
Diesing further demonstrates that presuppositional interpretations correlate closely with the VP-external scrambled position in German. The edge of the VP is marked by VP adverbs such as *immer* ‘always’; NPs which are scrambled outside of VP adverbs to this position cannot be extracted from. Compare, the sentences in (86) and (87) with a split *was für* (‘what kind’) interrogative. In (86), the NP is in its VP-internal position, as diagnosed by the fact that it is internal to the VP-adverb *immer*. Extraction of the *wh*- element *was* from this NP is grammatical. This is not the case, however, when the NP is external to the VP-adverb, as in (87). Here, subextraction is prohibited.

(86) Was<sub>i</sub> hat Hilda immer<sub>NP</sub>[ *t<sub>i</sub>* für Sonaten] gespielt?  
what has Hilda always for Sonatas played  
‘What kind of Sonatas did Hilda always play?’

(87) \*Was<sub>i</sub> hat Otto<sub>NP</sub>[ *t<sub>i</sub>* für Bücher] immer geschrieben?  
what has Otto for books always written

It appears that whatever is responsible for these effects in English and German could easily also be responsible for the clitic facts in Tagalog, although the implementation is not obvious at this point. The approach tentatively taken here relies on the notion that when possessor clitics are attached to specific NPs they modify a higher constituent and not the predicate head, as they do with existential and non-presupposed constituents. Possessor clitics must thus be visible not to the predicate head in such cases but rather to the referential index carried by the higher DP. Being

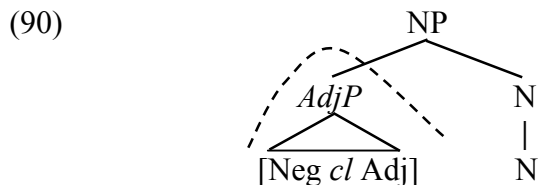
linearly contained within NP will render possessor clitics invisible to the higher DP bearing the referential index, as in (88)a. Positioning after the entire NP, on the other hand, as in (88)b, allows the clitic to be linearly contained by DP and thus visible to its referential index.



Note that some of the above cases of delayed positioning are not obviously NP constituents as several of these examples involved negation as the initial element. But as seen above, negation in Tagalog can also be NP-internal, i.e. constituent negation. It is precisely this structure which is posited for the examples seen earlier in (74) and (77)b. Constituent negation in conjunction with presuppositionality licenses the delayed clitic positioning witnessed above. A clearer case of constituent negation in DP is seen in (89), where we find the structure  $[[\text{NEG ADJ}] = \text{cl N}]$  for ‘their not good doings’. Placing the genitive clitic directly after negation violates the Clitic Visibility Condition, as is clear from the diagram in (90).

(89) bakà walà=ka=ŋ ka-málay~málay ay ma-sañkot  
 maybe NEG.EXT=2S.NOM=LNK EXT-INTNS~consciousness TOP PV.NVL-involve

{sa=hindi ma-ganda=**nila**=ŋ gáwá-in}  
 OBL=NEG ADJ-beauty=3P.GEN=LNK do-PV:NMLZ  
 ‘Maybe without realizing it, you will be involved in their not good doings.’<sup>43</sup>



<sup>43</sup> From: [www.filipinoes.net/top-4478.270.htm](http://www.filipinoes.net/top-4478.270.htm)



As expected, the possessor is outside the scope of negation when negation attaches directly to a daughter of NP, as in (89). When negation is clausal, both readings are possible, as shown by the minimally contrasting phrase in (91). This further suggests that the internal position of clitics is the unmarked one, allowing multiple readings.<sup>44</sup>

- (91) aŋ=hindi=**nila**=ŋ          ma-ganda=ŋ gáwá-in  
 NOM=NEG=3P.GEN=LNK ADJ-beauty do-PV:NMLZ  
 ‘the good doings which are not theirs / their not good doings’

It must be iterated that, regardless of the analysis chosen, the solution must be based on a soft constraint. If not, NP internal possessor clitics would be banned in nominative phrases, which are generally definite and presupposed. This is however far from the case, as can be seen for example by (92), in which the possessor clitic interrupts the definite nominative NP ‘my two ears’.

- (92) P<um>alapak=pò aŋ=ákiŋ          maŋa=kamay at paa  
 <AV.BEG>clap=POL NOM=1S.GEN PL=hand CONJ foot  
  
 pati=na          {aŋ=dalawa=**ko**=ŋ          tenga}  
 even=ALRD NOM=two=1S.GEN=LNK ear  
 ‘My hands and feet clapped, and even my two ears’<sup>45</sup>

In this subsection we have seen that referentiality/presuppositionality may effect clitic placement by licensing delayed positioning. This was explained by the requirement that clitics must be visible not to the predicate in such cases but rather to

---

<sup>44</sup> We can rule out the a priori plausible analysis of negation forming some type of syntactic constituent with the clitic to obtain the first reading in (91). As, shown in (i)a, these two elements cannot form an independent phrase on their own.

- (i) a. \*Hindi=nila                                  b. Hindi sa=kanila  
       NEG=3P.GEN                                NEG OBL=3P  
       ‘Not theirs’

<sup>45</sup> From: [www.sentidokomon.com/index.php?paged=4](http://www.sentidokomon.com/index.php?paged=4)

the dominating DP, which carries a referential index for the constituent. This goes some distance in explaining the asymmetry noted by Kroeger between X' constituents. Because referentiality is a property of (non-aspectual) NPs rather than (aspectual) VPs this asymmetry appears to target NPs as optionally impermeable constituents. As suggested here, this fact can be derived directly from the semantics rather than from categorial labels.

### 5.5.2 *Property predicates*

Based on our formulation of the Clitic Visibility Constraint, we find two basic property predicate constructions which should differ in their treatment of clitics based on their differing structures. In one of these, a preverbal adjectival phrase serves as the primary predicate followed by a subordinated verb phrase headed by an infinitive verb. This construction, shown in (93), can be analyzed as a clitic climbing environment, since only clitics may intervene between the adjectival predicate and the subordinated verb.

- (93) ma-bilis=ako=ŋ      ma-lasiŋ  
 ADJ-fast=1S.NOM=LNK PV.NVL-drunken  
 'I get drunk quickly.'

We expect that a complex adjectival phrase in this construction could be interrupted by clitics as it is this phrase which contains the predicate head. This prediction is borne out by data such as that in (94).

- (94) masyádo=(a)ko=ŋ      ma-bilis      mag-pa-lóko  
 ovebegy=1S.NOM=LNK ADJ-quick AV-CAU(REFL)-crazy  
 'I'm too quick to let people make me crazy.'<sup>46</sup>

---

<sup>46</sup> From: [http://ayka08.multiply.com/journal/item/5/the\\_change\\_in\\_me](http://ayka08.multiply.com/journal/item/5/the_change_in_me)

The second relevant construction contains a fronted adjectival phrase followed by a verb inflected for aspect. Here, the verb is set off from the preceding adjectival phrase by the =*ŋ/na* linker, and not the *∅/na* linker.

- (95) masyádo=*ŋ* ma-bilis na l<um>a~laŋoy si=*Kikoy*  
 overly=*LNK* ADJ-quick *LNK* <*BEG*>*INCM*~swim *P.NOM*=*K*.  
 ‘Kikoy swims too quickly.’

The =*ŋ/na* linker in this construction suggests that the two elements here are in a symmetric relationship and that, unlike in the former construction, there is no subordination here. If this is correct, we predict that a clitic associated with a predicate in the lower phrase cannot appear linearly embedded within a higher phrase as it would be rendered invisible to the predicate head. Preliminary investigation suggests that this prediction is correct: the adjective phrase internal positions for the clitic cluster in (96) have a reduced acceptability in comparison to the following external position.

- (96) Masyado[??=*ko*=*siya*]=*ŋ* ma-dalas[=*ko*=*siya*] na t<*in*>awag-an  
 overly=*1S.GEN*=*3S.NOM*=*LNK* ADJ-frequent *LNK* <*RL*>*call-LV*  
 ‘I called him too frequently.’

Note however that judgments are less clear when we are dealing with a single nominative clitic, as in (97), in such examples speakers appear to accept both positions equally.

- (97) masyado[=*siya*]=*ŋ* ma-dalas[=*siya*] na na-ki~kita  
 overly=*3S.NOM*=*LNK* ADJ-frequent *LNK* *PV.BEG.NVL*-*INCM*~call  
 ‘He was seen too frequently.’

This reveals a distinction between this construction and focus fronted obliques, which are judged categorically ungrammatical when interrupted by any (external)

pronominal clitics at all. The crucial difference seems to be in the fact that the adjectival construction at hand offers the additional option of treating the adjectival as the primary predicate with the following inflected verb as a type of relative modifier. That is, even when the verb is inflected, this construction may be roughly similar to that seen above in (94), in which the verb is in the infinitive and the adjective is part of the primary predicate.<sup>47</sup> However, a difference exists in that when the following verb is inflected and connected to the adjectival by the =*η/na* linker, it cannot assign case to clitic arguments in the higher phrase. Because clitics intervene between adjectivals and the following material in all cases considered here, the distinction only becomes salient when the adjectival predicate is complex. When the pronouns in question are associated with the lower predicate and must rely on it for case, they cannot appear embedded within the adjectival phrase. They can only appear in the embedded position if they get their case from the adjectival itself. This explains why it is only

---

<sup>47</sup> Note that there exists a difficulty with the idea that all adjectives in this construction may function as the main predicate of the clause rather than an adverbial modifier. While both manner and resultative type adverbs may appear as the initial element in the relevant construction, as in (i)a and b, respectively, only the former may take a subject when the verb is unambiguously contained within a subordinate clause. This is seen in (ii) where the verb is introduced within an if-clause, another common strategy for forming secondary predications in Tagalog. *tabiηi* ‘lopsided’ may serve as an independent matrix clause predicate but *masarap* ‘delicious’ cannot.

- (i) a.    *tabiηi=siya=η*                    *mag-lakad*  
           *lop\_sided=3S.NOM=LNK AV-walk*  
           ‘He walks lop-sidedly.’
- b.    *ma-sarap=siya=η*                    *mag-lútò*  
           *ADJ-delicious=3S.NOM=LNK AV-cook*  
           ‘He cooks deliciously.’
- (ii) a.    *tabiηi*    *si=Juan*    *kuη=mag-lakad*  
           *lop\_sided P.NOM=J. if=AV-walk*  
           ‘Juan walks lop-sidedly.’ (Lit. ‘He’s lop-sided if he walks’)
- b.    \**ma-sarap*    *si=Juan*    *kuη=mag-lutò*  
           *ADJ-delicious P.NOM=J. if=AV-cook*  
           (OK for ‘Juan is delicious when he cooks’)

nominative clitics which are licensed within complex adjectivals. Adjectivals are intransitive and unable to assign genitive case to an argument, as is obvious from the simpler contexts in (98).<sup>48</sup>

- (98) a. masyado=siya=ŋ      ma-dalas  
 overly=3S.NOM=LNK ADJ-frequent  
 ‘He was too frequent.’
- b. \*masyado=ko=siya=ŋ      ma-dalas  
 overly=1S.GEN=3S.NOM=LNK ADJ-frequent

### 5.5.3 Prepositional predicates

Prepositional predicates have the same basic syntactic properties of other intransitive predicates. The prepositional phrase, headed by the oblique marker *sa* or stative oblique *nása*, is in clause initial position and takes a following nominative marked subject. There are no copula elements in Philippine languages which differentiate verbal and non-verbal predication. Philippine languages in general are not rich in prepositions and it is commonly understood that prepositional notions in these languages are most often expressed by relational nouns, e.g., Tagalog *babà* ‘bottom’, *loob* ‘inside’, *kabilà* ‘other side’. These can be introduced as complements of the oblique marker *sa* and in turn take genitive complements to form what are functionally/translationally prepositional phrases, as shown in (99). As predicates,

---

<sup>48</sup> Counter-evidence to this claim is found in the attestation in (i), in which a genitive pronoun appears linearly embedded within an adjectival phrase followed by an inflected verb, for which it depends on for case. Whether or not this betrays a real possibility or is rather an error remains to be verified.

(i)      yuŋ                      *rhum coke*=kasi, masyádo=**ko**=ŋ      ma-bilis=‘ata <in>inom-Ø  
 that.NOM:LNK *rhum coke*=REAS overly=1S.GEN=LNK ADJ-quick=EPST <BEG>drink-PV  
 ‘because of that that rum-coke, I must have drunk it too quickly.’

oblique phrases are often headed by *nása* rather than *sa*, which indicates a stative (non-directional) meaning and is glossed here as STA.OBL.<sup>49</sup>

- (99) a. (ná)sa=ilálim    naᅇ=lamésa    b. (ná)sa=loob    naᅇ=kahon  
 (STA.)OBL=under GEN=table    (STA.)OBL=inside GEN=box  
 ‘underneath the table’    ‘inside the box’
- c. (ná)sa=labas    naᅇ=báhay    d. (ná)sa=ibábaw    naᅇ=dágat  
 (STA.)OBL=outside GEN=house    (STA.)OBL=surface GEN=sea  
 ‘outside the house’    ‘on the sea’

Some verbal elements have been grammaticalized with prepositional functions. These include *gáliᅇ*, *búhat* come\_ from ‘from’, *pa-punta* DIR-go ‘towards’ and *haᅇgaᅇ* limit(=LNK) ‘until’. These elements usually take oblique phrase complements, as shown in (100), but can also take bare and genitive marked complements as well. The verbal counterparts of the prepositions in (100) are shown in (101).

- (100) a. gáliᅇ sa=lunsod    b. pa-punta sa=ílog  
 from OBL=city    DIR-go OBL=river  
 ‘from the city’    ‘towards the river’
- (101) a. maᅇ-gáliᅇ sa=lunsod    b. p<um>unta sa=ílog  
 AV-from OBL=city    <AV>go OBL=river  
 ‘to come from the city’    ‘to go to the river’

Sityar (1989) and Kroeger (1996, 1998) claim that complex predicates headed by these elements optionally function as impenetrable constituents for the purposes of clitic placement. However, there is an overwhelming preference for them to host clitics directly and delayed positioning is often judged ungrammatical by speakers, in contradiction to Sityar’s original claim. Speakers’ judgments were corroborated by

<sup>49</sup> Oblique arguments cannot take *nása* while modifiers typically take *sa* but can also take *nása*. Oblique predicates take *sa* or *nása* depending on their directionality/stativeness.

comparing clitic placement in several common prepositional predicates using internet searches. Four of these examples are shown in (102) and (103), where the results of Google searches for each position are shown above the respective clitic position. The complement of a preposition can take either the oblique case, as shown in the (a) examples or the genitive case as in the (b) examples. The two fragments shown are based on searches for *galing sa/ng UP* ‘coming from UP’ (the University of the Philippines) and *galing sa/ng Manila* ‘coming from Manila’. (No examples involved focus fronting; in all cases the prepositional phrase constituted the main predicate.)

- (102) a. 20 hits                      0 hits  
           gálinj[=ako]    sa=UP[\*?=ako]  
           from=1S.NOM OBL=UP  
           ‘I’m from U.P.’
- b. 14 hits                      0 hits  
           gálinj[=ako]    naṅ=UP[\*?=ako]  
           from=1S.NOM GEN=UP  
           ‘I’m from U.P.’
- (103) a. 12 hits                      1 hit  
           gálinj[=ako]    sa=Mánilà [\*?=ako]  
           from=1S.NOM OBL=MANILA  
           ‘I’m from Manila.’
- b. 15 hits                      0 hits  
           gálinj[=ako]    naṅ=Mánilà[\*?=ako]  
           from=1S.NOM GEN=MANILA  
           ‘I’m from Manila.’

As can be seen, out of 62 instances, only 1 attestation was found for phrase final positioning. Taken together with the fact that speakers judge such positioning as seriously degraded, this belies the claim that a higher degree of optionality exists for clitic positioning in non-verbal predicates than for verbal predicates in Tagalog.

The situation, however, is markedly different with *hanǵanǵ* ‘until’. In contrast to the majority of prepositional elements, *hanǵanǵ* ‘until’ never allows 2P clitics to follow directly, regardless of syntactic context, as shown in (104).

- (104) hanǵanǵ[\*=pa=ako] sa=liŋgo[=pa=ako]  
 until=still=1S.NOM OBL=Sunday  
 ‘I’m until Sunday.’

*Hanǵanǵ* also does not license extraction of its complement. The difference between *hanǵanǵ* and *ǵalıŋ* in this respect is seen in (105), in which an oblique interrogative element must strand the prepositional *ǵalıŋ* in (105)a but cannot do so with *hanǵanǵ* in (105)b. In the grammatical version of (105)b, shown in (105)c, the interrogative element pied-pipes the preposition and the entire constituent acts as a clitic host.

- (105) a. saan=ka ǵalıŋ?  
 where=2S.NOM from  
 ‘Where are you from?’  
 b. \*kailan[=ka] hanǵanǵ[=ka]?  
 when=2S.NOM until  
 ‘When are you (at X, doing X, etc.)?’  
 c. hanǵanǵ kailan=ka?  
 until when=2S.NOM  
 ‘Until when are you (at X, doing X, etc.)?’

This situation has an analogue in Serbo-Croatian, as noted by Progovac (1996) (see also Wilder and Čavar 1994, Bošković 2000, 2001, Zec 2005). This is shown in (106), where the auxiliary clitic must follow the entire prepositional phrase rather than just the preposition. What makes this case interesting is that the preposition is clearly a stressable prosodic word in Serbo-Croatian. This datum has been used to argue against prosodic inversion type analyses in which the clitic inverts with a following prosodic word regardless of its syntactic status. The fact that prepositions are also ineligible for



movement is seen by proponents of a syntactic approach to indicate that clitic hosts arrive in their position via ordinary syntactic movement.

- (106) a. Prema[\*=su] Mileni[=su] Milan i Jovan išli *Serbo-Croatian*  
 toward=AUX Mileni Milan and Jovan walked  
 ‘Milan and Jovan walked toward Milena’

Anderson (2005:118) shows, however, that any approach which can demand contiguity between certain constituents will also be able to handle both the clitic placement facts and the movement facts in a unified manner. We follow Anderson here treating the impenetrability of these types of phrases as resulting from the constituency of certain prepositions and their complements. Again, the fact that neither argument nor adverbial clitics are able to penetrate the *hangaj*+X constituent shows that neither domains nor visibility is at stake here.

#### 5.5.3.1 *Complex prepositional predicates*

If prepositions were on par with verbs and nouns and took NP complements, we would expect that this complement would be impenetrable for argument clitics, as such clitics would be in an embedded domain which blocked their visibility to the predicate head. This is however not the case, as pronominal clitics typically interrupt the nominal complement of an oblique phrase as shown from the attestations in (107)-(108).

- (107) nása=ma-laki=**ka**=ŋ paŋánib, Abraham  
 STA.OBL=ADJ-big=2S.NOM=LNK danger Abraham  
 ‘You are in great danger, Abraham.’

- (108) nása=ma-ganda=**sila**=ŋ kulujan  
 STA.OBL=ADJ-beauty=3P.NOM=LNK cage

‘They’re in a beautiful cage.’<sup>50</sup>

There is good evidence that prepositions are not like other lexical categories in that they are not lexical but rather functional (van Riemsdijk 1990, Grimshaw 1991). Grimshaw (1991) and Emonds (1985), for instance, treat P as parallel to complementizers in clausal projections. Baker (2005), in his theory of lexical categories, explicitly treats prepositions as non-predicational, accounting for their well-known inability to function as bare predicates without the help of copulas or other functional material. Baker further shows that in Edo, a West African language, prepositional phrases cannot even form predicates with the help of the predicational element used for nominal predicates *yé/rè*, as seen in (109)a. Locative and directional type predicates must rather be expressed with locative or posture verbs, as in (109)b and c (Baker 2005:314).

- |   |  |   |
|---|--|---|
| (109)a. *òzó (yé/rè) vbè òwá<br>Ozo PRED at house<br>(*Pred+PP) | b. òzó rré òwá<br>Ozo is.at house<br>‘Ozo is in the house.’<br>(locative verb) | c. òzó mùdiá yè esuku<br>Ozo stand at school<br>‘Ozo is at school.’<br>(posture verb) |
|---|--|---|

Although Austronesian languages typically have no copular elements or clear candidates for Pred<sup>0</sup>, we can see a Tagalog parallel in the fact that there exists no bare preposition which can function as a predicate, i.e., as in English “She’s in”. All such sentences must be expressed by relational nouns. Baker (2005:315) draws the following conclusion: “These data suggest that PPs do not license specifiers directly. Indeed, it seems that not even a functional head like Pred is enough to create a specifier for a PP in most languages. The reason for this is presumably semantic rather than syntactic: we can say that there is no theme role implicit in the lexical meaning of the adpositions that a Pred can bring out.” If Baker is correct in laying the fault on

<sup>50</sup> From: <http://www.sabong.net.ph/forum/showthread.php?t=27371>



types of structures which have important implications for a theory of clisis: coordination and ellipsis.

## 5.6 Coordination

Coordination of predicates has been claimed to only allow clitic placement at the right edge of larger coordinated structure, as shown in (112) (S&O:187-189, Kroeger 1993:121).<sup>51</sup>

- (112) k<um>a~kain[??=ako] at <um>i~inom[=ako]  
<AV>INCM~eat=IS.NOM CONJ <AV>INCM  
'I am eating and drinking.'

Similar facts in Serbo-Croatian have been used to argue for the syntactic nature of 2P (Cavar 1999). Because extraction out of conjoined structures is also generally forbidden (Ross 1967), it is argued that movement within the syntactic component is generally employed to position the clitic host to the left of clitics. When such movement fails for syntactic reasons, it is expected that the constituent in question cannot function as a clitic host. But as has been repeatedly emphasized here, this cannot be the correct approach to Austronesian clitics as general movement is far more restrictive than one would be led to believe on the basis of what can serve as a host for 2P clitics. If extraction facts are to be unified with clitic placement facts then the connection cannot be a unidirectional one from narrow syntax to clitic syntax.

The first question to ask within the present framework is whether the problem is due to a violation of the Clitic Visibility Condition or is rather morphological in

---

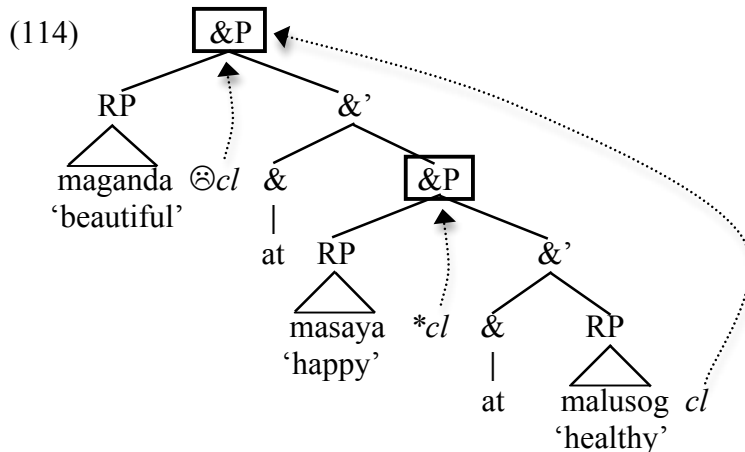
<sup>51</sup> Placement of the clitic after the first constituent is not strictly speaking ungrammatical but it does not entail a predicate coordination reading. Rather, the second predicate must be interpreted as having a null subject. The awkwardness which ensues is due to the fact that pro-drop is not easily licensed with first and second person subjects, thus forcing an anaphoric relation which ignores the immediately preceding pronoun.

nature. As above, this can be easily diagnosed by testing for the possibility of adverbial clitics in the same positions. As seen in (113), adverbial clitics are perfectly grammatical in inner conjuncts.

(113) ma-laki=**na=pala**, ma-lusog=**pa=yatà** at tiyak na murà  
 ADJ-big=ALRDY=MIR ADJ-healthy=STLL=EVID CONJ certainly LNK cheap

aŋ=maŋa=manok=nila  
 NOM=PL=chicken=3P.GEN  
 ‘Their chickens are surprisingly already big, probably even healthy, and certainly cheap.’

It is possible to conclude from this that, to the extent argument clitics are awkward within inner conjuncts, the problem is one of visibility. Taking an asymmetric analysis of coordination, we can posit the (pruned and abbreviated) structure in (114) as a typical conjunction. Applying the clitic visibility condition to such a structure predicts that a clitic after the first or last conjunct should be visible to all of the conjoined predicate heads, as shown. While visibility correctly rules out intermediate positions for argument clitics, it appears to overgenerate in allowing placement after the first conjunct (i.e., the clitic marked by ☹). If both initial and final positions pass the visibility condition we expect that the leftward alignment constraint would rule out the final position, but this is clearly not the case.

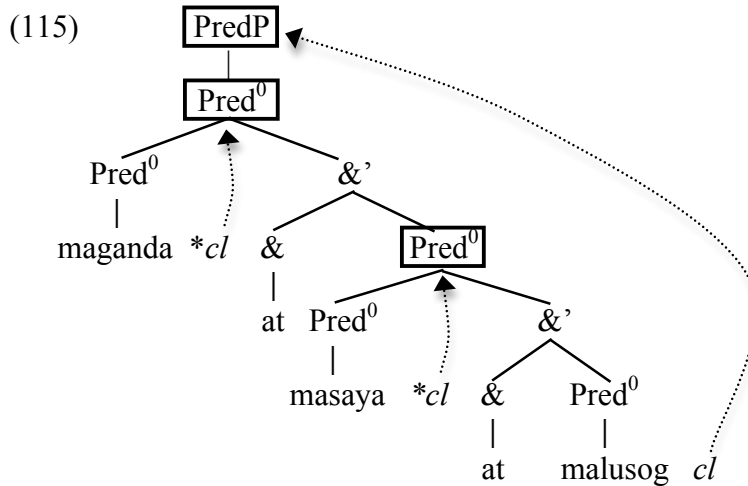


There thus appears to be a discrepancy between the predictions of the current theory and the reported facts.<sup>52</sup> But until now we have been assuming that the visibility condition applies to argument clitics and the terminal node hosting the predicate head which assigns them thematic roles. Another possibility is that the relevant category is larger than Pred<sup>0</sup>, i.e., PredP, or other such functional projections associated with argument structure. In this case, conjunction of Pred<sup>0</sup> would yield a different output from conjunction of the larger functional category. In the former case, internal clitics would be invisible to the higher functional projection and in the latter case they would not be. The visibility status of clitics in head conjunction constructions is shown in (115). Internal positions are invisible to the phrasal node and may thus be ruled out. The optionality of conjoining heads, as in (115), or larger categories, as in (114), correctly predicts that non-peripheral positions with more than two conjuncts are always ungrammatical for pronominals (although not for adverbials which need not satisfy visibility, as seen in (113) above). This is because spelling out clitics in medial positions will always guarantee invisibility to both the higher conjunct head and the larger containing category.

---

<sup>52</sup> Attestations of cliticization to inner adjuncts, as in (i) and (ii), do not appear to be terribly uncommon. Nonetheless, all speakers accept positioning clitics after the entire conjoined constituent and this is taken to be the unmarked position. Kroeger (1993:121 fn.7) suggests that examples like these involve prodrop of the second conjunct, which can be diagnosed by an interpretation where the subject of the second conjunct does not corefer with that of the first one.

- (i) ma-ganda=siya      at      mukha=ŋ ma-bait!  
 ADJ-beauty=3S.NOM CONJ face=LNK ADJ-nice  
 ‘She is beautiful and seems nice!’ (From: [www.voy.com/212232/](http://www.voy.com/212232/) )
- (ii) ma-ganda=**ako**      at      ma-talino at      ma-bait  
 ADJ-beauty=1S.NOM CONJ ADJ-smart CONJ ADJ-nice  
 ‘I’m beautiful and smart and nice.’  
 (From: <http://www.witchmag.com.ph/board/index.php?topic=21225.270>)



This approach makes another important prediction: if final positioning with conjunction is the result of conjoining smaller categories, we should be able to eliminate certain positioning possibilities when there is overt evidence for more structure, e.g., phrasal arguments. In particular, we expect that conjunction of larger categories should disallow final positioning and require placing the clitic after the first available host within the first conjunct. This prediction is borne out, as seen below from (116) and speakers' rejection of final positioning in the similar (117).

(116) k<um>ain=ako      nan=pizza at    <um>inom    nan=diet rootbeer  
 <AV.BEG>eat=1S.NOM GEN=pizza CONJ <AV.BEG>drink GEN=diet\_root\_beer  
 'I ate pizza and drank diet root beer.'<sup>53</sup>

(117) ?\*k<um>ain    nan=pizza at    <um>inom=ako  
 <AV.BEG>eat GEN=pizza CONJ <AV.BEG>drink=1S.NOM

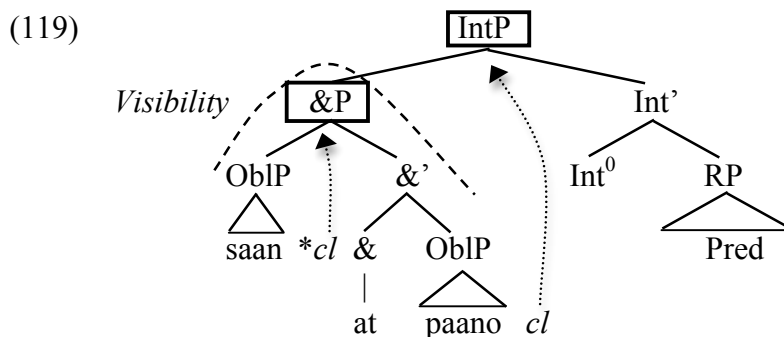
If conjunction of phrases is possible, as in (116), we must ask why this option is not available when there is no overt object, as in (112). This may be no more than a simple economy constraint on conjunction: only include as much functional structure as there is overt evidence for. In the case of bare verbs, economy may demand for

<sup>53</sup> From: [www.yehey.com/boards/Default.aspx?g=posts&t=142353](http://www.yehey.com/boards/Default.aspx?g=posts&t=142353)

some speakers that conjunction be of the minimal  $V^0$ . For other speakers, no such economy condition comes into play. Note that the predictions are more categorical for conjunction of hosts outside the predicate phrase. For instance, there is no possibility of argument clitics interrupting conjoined interrogatives, as in the attested example in (118).

- (118) Ito=ba=ŋ      si=Santiago Cid at      si=Vicente Pons ay alam  
 this=QM=LNK P.NOM=S.C.      CONJ P.NOM=V.P.      TOP know  
 ‘This Santigao Cid and Vicente Pons, did (they) know  
  
 kuŋ=saan      at      paano=**ninyo**      na-kúha-∅      aŋ=jeep?  
 COMP=where CONJ how=2P.GEN      NVL.BEG-get-PV NOM=jeep  
 where and how you obtained the jeep?’<sup>54</sup>

This is because, in this case, regardless of the size of the conjuncts, an argument clitic sandwiched by material in [Spec,IntP] would be invisible to the predicate, as illustrated in (119). Just as with complex phrasal hosts in this position, the clitic must follow all material in [Spec,IntP] to be visible to a lower predicate.



Again, just as with complex phrases in [Spec,IntP], we expect to see an asymmetry between argument and adverbial clitics and this is what we find. When we add the question marker to the example in (118), as shown in (120), the asymmetry is

<sup>54</sup> From: [http://www.lawphil.net/judjuris/juri1997/feb1997/gr\\_98252\\_1997.html](http://www.lawphil.net/judjuris/juri1997/feb1997/gr_98252_1997.html)



revealed. While the pronominal clitic must follow the entire conjoined phrase, the question marker can follow the first conjunct directly. Again, this is because adverbial clitics are not dependent on any overt material in the clause for their interpretation and thus not bound by the Clitic Visibility Condition. Their placement, is therefore purely prosodic. This difference will also play an important role in the following section.

- (120) Saan[=**ba**][\*=**ninyo**] at paano[=**ba**][=**ninyo**] na-kúha-Ø aŋ=*jeep*?  
 where=QM=2P.GEN CONJ how NVL.BEG-get-PV NOM=*jeep*  
 ‘Where and how did you get the jeep?’

In this subsection we have explained several facts about clitic positioning in conjoined structures and seen that clitics must in fact be visible to a functional category higher than the minimal Pred<sup>0</sup>.

## 5.7 Ellipsis

The Clitic Visibility Condition requires that clitics be in a certain syntactic configuration in relation to the predicate it is associated with but it has not yet been discussed at which level this relation must hold. In particular, we have not yet seen any evidence for or against the possibility that this is entirely a PF relation which can be bled by operations such as deletion. As it turns out, it is the case that deletion can bleed the satisfaction of the Clitic Visibility Condition. This perhaps can be predicted from the fact that clitics are not in their surface position in the underlying structure. Interestingly, ellipsis further validates the distinction drawn here between argument and adverbial clitics, specifically, the notion that argument clitics must be visible to certain projections while adverbial clitics, being purely functional, have no such structure to relate to.

Richards (2003) investigates ellipsis in Tagalog and uses evidence from clitics to argue for the size of the deleted constituent in three kinds of ellipsis. Richards associates all clitics with syntactic projections and treats their inability to appear with certain types of ellipsis as the result of deletion of those projections. Because his assumptions regarding the nature of clitics in Tagalog are significantly different than the view argued for here, his proposal requires some background.

Based on certain facts about strict versus sloppy readings with apparent DP-ellipsis, Richards argues that what appears to be DP-ellipsis in Tagalog is in fact vP ellipsis (cf. Otani & Whitman 1991 for Japanese). This ellipsis is claimed to exclude pronominal clitics, as shown by (121)a and b, where the pronouns *ako* 1S.NOM and *siya* 3S.NOM obligatorily surface despite deletion. In (121)c, however, we see that when the antecedent is a proper name, a following subject is deletable. Richards observes that, “just when the subject of the first conjunct is a clitic pronoun, the subject of the second conjunct cannot be elided by vP-ellipsis”. The necessary assumption (Richards 2003:fn.5) is that deletion is licensed in (121)c because what is being deleted there is not a clitic but rather a full NP.<sup>55</sup>

- (121) a. S<in>abi-Ø=ko=ŋ            mag-bí~bigay=ako  
           <BEG>say-PV=1S.GEN=LNK AV-INCM~give=1S.NOM  
           GEN=money OBL=church at    nag-bigay=ŋà=\*(ako)            [\_\_\_\_\_]  
           naŋ=pera    sa=simbahan CONJ AV.BEG-give=EMPH=1S.NOM  
           ‘I said I would give money to the church, and I did.’
- b. S<in>abi-Ø ni=Juan na    mag-bi~bigay=siya  
           <BEG>say-PV P.GEN=J. LNK AV-INCM~give=3S.NOM

---

<sup>55</sup> Presumably, for binding theoretic reasons which are not made explicit in the paper, insertion and deletion of a proper name in the second conjunct is only possible when the immediate antecedent is a proper name and disallowed when it is a pronominal.

naŋ=pera sa=simbahan at nag-bigay=ŋà=\*(siya) [\_\_\_\_\_]
   
GEN=money OBL=church CONJ AV.BEG-give=EMPH=3S.NOM
   
'Juan said he would give money to the church, and he did.'

- c. S<in>abi-∅=ko=ŋ mag-bi~bigay si=Juan
   
<BEG>say-PV=1S.GEN=LNK AV-INCM~give P.NOM=J.

naŋ=pera sa=simbahan at nag-bigay=ŋà=(siya) [\_\_\_\_\_]
   
GEN=money OBL=church CONJ AV.BEG-give=EMPH=3S.NOM
   
'I said Juan would give money to the church, and he did.'

Next, Richards argues for a distinction between pronominal and adverbial clitics in complement-of-Neg ellipsis, a deletion pattern which is exemplified for Tagalog in (122).

- (122) Hindi=ko alam kuŋ nag-bigay si=Juan naŋ=pera sa=simbahan,
   
NEG=1S.GEN know if AV.BEG-give P.NOM=J. GEN=money OBL=church

pero s<in>abi-∅ ni=Maria na hindi [\_\_\_\_\_]
   
but <BEG>say-PV P.GEN=M. LNK NEG
   
'I don't know if Juan gave money to the church, but Maria said he didn't'

Here, unlike in vP-ellipsis, pronominal clitics are not licensed at all, as can be seen in (123)a. Interestingly, however, adverbial clitics are completely grammatical in the same context, as shown by (123)b, where the adverbials *pa*, *daw* and *yatà* are all shown to be possible after negation.

- (123) Hindi=ko alam kuŋ nag-bigay=ako naŋ=pera sa=simbahan,
   
NEG=1S.GEN know if AV.BEG-give=1S.NOM GEN=money OBL=church
   
'I don't know if I gave money to the church...'

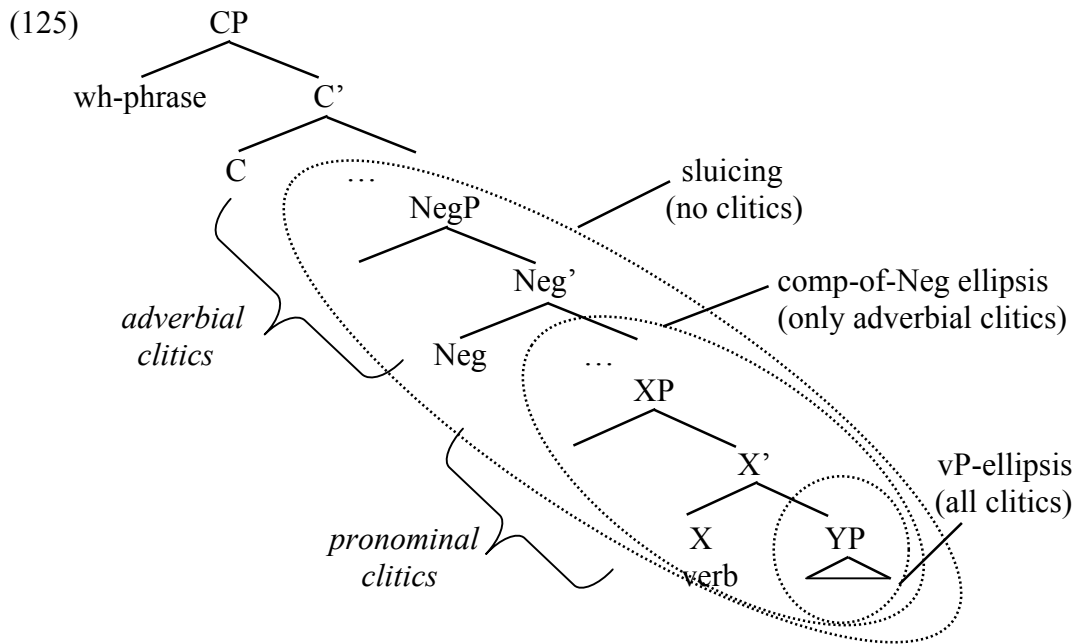
a. pero s<in>abi-∅ ni=Maria na hindi(=\*ako)
   
but <BEG>say-PV P.GEN=M. LNK NEG
   
'but Maria said I didn't'

b. pero s<in>abi-∅ ni=Maria na hindi=pa/=raw/=yatà
   
but <RL>say-PV P.GEN=M. LNK NEG=STILL=RPRT=APRNT
   
'but Maria said I didn't yet/reportedly/apparently' (Richards 2003:235)

Finally, Richards argues that in a third type of ellipsis, sluicing, neither pronominal nor adverbial clitics are licensed. In (124)a, Richards shows that insertion of a pronominal clitic subject is ungrammatical and in (124)b that adverbial clitics are likewise ungrammatical. This is claimed to contrast with (124)c, in which a full continuation allows adverbial clitics and necessitates a pronominal clitic (if the subject happens to be pronominal).

- (124)        Gusto=ko=ŋ        b<um>alik sa=Pilipinas,  
                  want=1S.GEN=LNK <AV>return OBL=Philippines  
                  ‘I want to back to the Philippines...’
- a.        pero hindi=ko        alam kuŋ        kailan(\*=ako)  
                  but NEG=1S.GEN know COMP when=1S.NOM  
                  ‘but I don’t know when.’
- b.        \*pero hindi=ko        alam kuŋ        kailan=ŋà /≠kayà /≠naman  
                  but NEG=1S.GEN know COMP when=EMPH =SPEC =SWTCH  
                  ‘but I don’t know when indeed/I wonder/by contrast.’
- c.        pero hindi=ko        alam  
                  but NEG=1S.GEN know
- kuŋ        kailan=ŋà/≠kayà/≠naman \*(=ako)        Ø-bá~balik  
                  COMP when=EMPH =SPEC =SWTCH 1S.NOM AV-INCM~return  
                  ‘but I don’t know when indeed/I wonder/by contrast I will return.’  
                  (Richards 2003:235)

Richards (2003:236-7) summarizes his analysis of the three types of ellipsis in Tagalog with their correlating consequences for clitics in the tree in (125). The circled portions represent the targets of sluicing, comp-of-Neg ellipsis and vP-ellipsis. He concludes from the clitic facts that adverbial clitics are generated higher in the tree between C and Neg while pronominal clitics are generated lower between Neg and the raised position of the verb.



There are several empirical difficulties with this analysis mostly centering around certain claims of (un-)grammaticality in the above data. First of all, the judgments reported by Richards in (121) give the impression that there exists a robust difference between clitic deletion and full-NP deletion, with only the latter being licensed by  $vP$ -deletion. Unfortunately, native speakers with whom I consulted could not substantiate any distinction between (121)b and c. Rather, as discussed by Himmelmann (1999), pro-drop is a highly restricted phenomenon in Tagalog, only applying to third person arguments under particular circumstances. Thus, although the more natural continuation in both (121)b and c includes the pronominal subject, both examples license pro-drop of this argument in contrast to sentences with a first or second person subject, as in (121)a, which do not. The sluicing data is also problematic. The crucial example is in (124)c, which is meant to show that non-deletion of the complement of C in ellipsis licenses both adverbial and pronominal clitics. However, this example is not well constructed and as a result, somewhat misleading. As can be expected, adverbial clitics interact in a complex manner with

the illocutionary force and modality of the larger sentential context. The main problem with (124)c is that it is embedded under a desiderative predicate *gusto* ‘want’, which makes it awkward even without additional clitics. The natural reading of this – and in fact, the only expected reading under standard theories of sluicing, e.g., Merchant (2001) – is one in which the desiderative predicate is also interpreted as part of the deleted material, i.e., as in (126). The sentence must thus be interpreted such that the speaker doesn’t know his or her own desires, an arguably unusual state of affairs.

- (126) ?*gusto=ko=ŋ*      *b<um>alik*    *sa=Pilipinas*,  
           want=1S.GEN=LNK    <AV>return    OBL=Philippines  
  
           *pero hindi=ko*    *alam* *kun*    *kailan* ~~*gusto=ko=ŋ*~~ ~~*b<um>alik*~~ ~~*sa=Pilipinas*~~  
           but    NEG=1S.GEN    know    COMP    when  
           ‘?I want to go back to the Phil., but I don’t know  
           when I want to go back to the Phil.’

Further adding the speculative adverbial clitic *kayà*, emphatic *ŋà* or switch topic *naman* to this sentence only compounds the problem by emphasizing the speakers ignorance. Now note that the final conjunct in (124)c, the grammatical example containing adverbial clitics without deletion, is not equivalent to the deleted portion in (124)b, the sluiced sentence. Crucially, the continuation offered by Richards does not contain the desiderative predicate which triggers the infelicitous reading. The reason that speakers react to (124)c more favorably thus has nothing to do with the adverbial clitics putatively being licensed by a portion of phrase structure not present in the sluicing example.

In fact, sluicing and sentential fragments are completely felicitous with adverbial clitics so long as other semantic factors are controlled for.<sup>56</sup> For instance, the sentence in (127) represents an absolutely commonplace construction.

- (127) Ø-pu~punta=sila sa=Amerika, pero kailan=kayà [\_\_\_\_]?  
 AV-INCM~go=3P.NOM OBL=A. but when=SPEC  
 ‘They’re going to America, but when, I wonder?’

Three attestations of classic sluicing constructions containing adverbial clitics are given in (128)-(130). In each case, the presence of adverbial clitics is considered unremarkable by speakers.

- (128) Naka~ka-irita=daw, sabi naŋ=kapatid=ko,  
 AV.BEG~INCM-irritate=RPRT say GEN=sibling=1S.GEN

pero hindi=ko alam kuŋ bakit=**naman**.  
 but NEG=1S.GEN know COMP why=SWTCH  
 ‘It’s allegedly irritating, my sibling says, but I don’t know why.’<sup>57</sup>

- (129) D<in>ala-Ø naŋ=*mama*=ko at naŋ=*aunt*=ko sa=hospital  
 <BEG>take-PV GEN=mom=1S.GEN and GEN=aunt=1S.GEN OBL=hospital  
 péro di=daw=nila t<in>aŋgap-Ø.  
 but NEG=RPRT=3P.GEN <BEG>accept-PV

sábi naŋ=*sis*=ko di=niya alam kuŋ bakit=**daw**  
 say GEN=sister=1S.GEN NEG=3S.GEN know COMP why=RPRT  
 ‘My mom and my aunt took (them) to the hospital but they didn’t accept (them). My sister said she didn’t know why (reportedly).’<sup>58</sup>

<sup>56</sup> Merchant (2001:107 fn.12) comments that fragments obey the same principles as sluicing and therefore can be subsumed under the same PF-deletion analysis.

<sup>57</sup> From: <http://www.thedigitalpinoy.com/thread/11/1236/3>

<sup>58</sup> From: [www.sendami.com/2007\\_01\\_01\\_archive.html](http://www.sendami.com/2007_01_01_archive.html)

(130) siguro=naman ma-gi~gin gold=din yuŋ kay=sheryn,  
 probably=SWTCH AV-INCM~become gold=also that:LNK P.OBL=S.

kayà\_laŋ, we never know=pa kuŋ kailan=**kayà**?  
 just\_that, we never know=still COMP when=SPEC  
 ‘Sheryn’s will probably also become gold, it’s just that,  
 we’ll never know when.’<sup>59</sup>

To review, we saw that in the putative vP-ellipsis construction in (121), where the predicate head is present, argument clitics are licensed. In comp-of-Neg ellipsis, seen in (123) above, the predicate head is deleted and argument clitics are not licensed. Finally, in sluicing, the predicate head is deleted and again argument clitics are not licensed. In all three ellipsis types, it was seen that adverbial clitics are in fact licensed. Consequently, the true generalization appears to be that adverbial clitics may surface wherever they are semantically licensed whereas argument clitics must satisfy the visibility condition on the surface.<sup>60</sup> Deletion of the predicate head with which a pronominal clitic is associated deprives it of the chance to satisfy the visibility condition and thus ensues in ungrammaticality.<sup>61</sup>

<sup>59</sup> From: <http://www.titikpilipino.com/news/index.php?aid=213&offset=160&expanded=1>

<sup>60</sup> Incidentally, Richards’ claim for adverbial clitics in sluicing constructions does appear correct for Serbo-Croatian. Bošković (2001:34) shows that the question marker *li* does not survive in sluicing constructions such as (i). I have no explanation for this variation except for the possibility that Serbo-Croatian *li* may be in licensing relationship with a deleted functional projection.

- (i) Vidi nekoga. \*Koga li vidi?  
 sees somebody who QM sees  
 (For, ‘He sees somebody. Who?’)

<sup>61</sup> Note that deletion need not be crucially ordered in relation to the Clitic Visibility Condition, rather the clitic visibility condition is a purely surface constraint and cannot be ordered before any syntactic operation, whether in narrow syntax or PF.

Bošković (2001:71) makes the same empirical observation for Serbo-Croatian (i), but draws from it a rather different conclusion, namely, that 2P is a requirement of the phonology which can be circumvented by deletion. But this is predicated on the presence of clitics in argument position, an unlikely prospect for Philippine languages as argued here.

- (i) Marija=ga nije poljubila, a Ana jeste poljubila-ga  
 Maria=him not.is kissed and Ana IS kissed him  
 ‘Maria didn’t kiss him, but Ana did.’



## 5.8 Compounding, reduplication and clitic placement

Until now we have been assuming a working definition of the word for the purposes of clitic placement. In this section we take a closer look into morphological word constituency looking in particular at several reduplication processes.

The criteria for hosting clitics includes being a prosodic word and a maximal morphological word. Elements which are maximal morphological words but not prosodic words include case markers and other monosyllabic functional heads. Their inability to host clitics is derived from \*WEAKSTART and needs no further comment. On the other hand, we find several morphological processes, chiefly compounding and reduplication, which yield constituents that satisfy prosodic wordhood but are not maximal morphological words.

The linker, an element found in almost all Philippine languages, intervenes between elements in a modificational relationship, functioning also as a relative marker as well as a complementizer. The linker serves as a sufficient, but not necessary, indicator of morphological wordhood. Adverbial modifiers such as *lági* ‘always’ are typically attached to the elements they modify via the linker, as seen in (131)a, but negation, also an independent word in Tagalog, is not, as seen in (131)b.

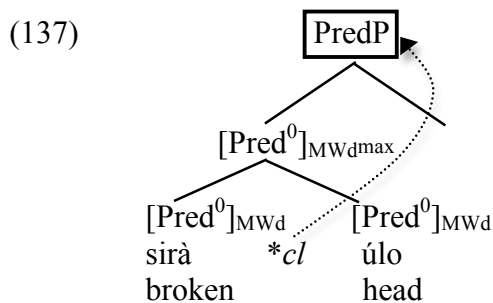
- (131) a. *lági=ŋ ma-saya*                      b. *hindi ma-saya*  
          always=LNK ADJ-happy                NEG ADJ-happy  
          ‘Always happy.’                        ‘not happy’

The linker’s distribution with 2P clitics is noteworthy: 2P clitics must appear before the linker and can never follow it, as seen in (132). Note also that the presence of the linker is wholly conditioned by the clitic host and not the clitic. As may be expected, the linker never appears connecting clitics to each other.



- (135)a. báhay-bato  
house stone  
'stone house'
- b. Báhay[\*=siya(=ŋ)] bato[=siya]  
house=3S.NOM=LNK stone  
'It's a stone house.'
- (136)a. sirà-úlo  
broken-head  
'crazy'
- b. sirà[\*=ka(=ŋ)] úlo[=ka]  
broken=2S.NOM=LNK head=2S.NOM  
'You're crazy.'

These examples show clearly that clitic placement is sensitive to morphological structure as well as to prosodic structure. This is in fact already predicted by the Clitic Visibility Condition. It was seen in §5.6 on the basis of clitic placement with coordinated predicate heads that the Clitic Visibility Condition must require visibility to the phrasal constituent containing the predicate rather than the predicate head itself. It thus follows that being linearly contained within a single morphological constituent will render a clitic invisible to structure outside that maximal word, as shown in (137),



Just as with compounds, the output of morphological doubling operations on property-type predicates cannot be intruded upon by clitics. Different types of

reduplication in Tagalog operate on different morphological categories (Carrier-Duncan 1978), which makes it an excellent diagnostic for morphological constraints on clitic placement. We examine three types of reduplication in Tagalog in this section and observe the corresponding clitic placement patterns.

In moderative reduplication, the first foot of a property-type root is reduplicated without the intervention of the linker, e.g. *ma-ganda* ‘beautiful’ → *ma-ganda-ganda* ‘somewhat beautiful’. In intensive reduplication, the entire adjective is reduplicated with the intervention of the linker, e.g., *ma-ganda* ‘beautiful’ → *ma-ganda=η ma-ganda* ‘extremely beautiful’. Two facts suggest that, unlike moderative reduplication, the target of intensive reduplication is an entire morphological word: (i) it requires the linker, which is never found word internally, (ii) it does not abide by the maximality constraint on word-internal reduplicants which imposes a disyllabic ceiling. This latter point can be seen in the different outputs of the two types of reduplication on a trisyllabic root, shown in (138).

- (138)a. *ma-bala~balahíbo*  
 ADJ-MDRT-body.hair  
 ‘somewhat hairy’
- b. *ma-balahíbo=η~ma-balahíbo*  
 ADJ-body.hair=LNK~ADJ-body.hair  
 ‘very hairy’

Despite the difference in morphological domains, clitics cannot render the reduplicant non-contiguous to the base in either moderative or intensive reduplication, as shown in (139)-(140) (Kroeger 1993:121).

- (139) *ma-ganda[\*=siya]~ganda[=siya]*  
 ADJ-beauty=3S.NOM~beauty=3S.NOM  
 ‘She is somewhat beautiful.’
- (140) *ma-ganda[\*=siya]=η~ma-ganda[=siya]*  
 ADJ-beauty=3S.NOM=LNK~ADJ-beauty=3S.NOM  
 ‘She is extremely beautiful.’

This is expected trivially for moderate reduplication as the reduplicant is smaller than a morphological word. Clitic placement with intensive reduplication, on the other hand, requires the Clitic Visibility Condition since it clearly constitutes two morphological words. Given this condition, however, the expectations are clearly similar to that of compounding. Intensive reduplication is a word level morphological process which produces larger words and clitics trapped between two parts of a single morphological entity will not be visible to the predicate phrase as required.

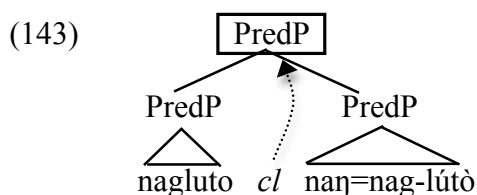
The third type of reduplication examined here is iterative reduplication. This type of reduplication appears to double aspectual predicate heads, adding iterative semantics. Crucially, the reduplicant is embedded under a genitive case marker, as shown in (141).

- (141) nag-lútò      naŋ=nag-lútò      naŋ=pansit      aŋ=taga-pag-lútò  
 AV.BEG-cook   GEN=AV.BEG-cook   GEN=noodles   NOM=PROF-TR-cook  
 ‘The cook cooked noodles continuously.’

The target here appears to be the predicate head *nag-lútò* ‘cooked’ but as is clear from the fact that Tagalog case markers are phrase level clitics rather than word level affixes, we are no longer dealing here with word level categories but rather phrasal ones. Nonetheless, the phrasal domain does not include objects or modifiers, as seen by the ungrammatical (142). We can thus understand the target of reduplication as the minimal phrase hosting the predicate, that is PredP. Objects are either generated in a lower phrase or have evacuate PredP by the time iterative reduplication applies.

- (142) \*nag-lútò      naŋ=pansit      naŋ=nag-lútò      naŋ=pansit  
 AV.BEG-cook   GEN=noodles   GEN=AV.BEG-cook   GEN=noodles

Because iterative reduplication targets a phrasal projection and it has been argued above that pronominal clitics must be visible to the minimal phrasal projection containing the predicate we expect intervention by clitic to be possible and this is what we find. As shown by (144), clitics not only can but must intervene between the base and the reduplicant in this construction.



(144) nag-lútò[=sila]      naŋ=nag-lútò[\*=sila]      naŋ=pansit  
 AV.BEG-cook=3P.NOM      GEN=AV.BEG-cook      GEN=noodles  
 ‘They cooked noodles continuously.’

We have seen in this brief section that the Clitic Visibility Condition perfectly predicts the interruptability of different morphological constructions in Tagalog. Intrusion by clitics is always licensed if the clitics can maintain their surface relationship with the predicate phrase.

## 5.9 Alternative theories of Tagalog clitic syntax

### 5.9.1 Kroeger 1993: the first daughter approach

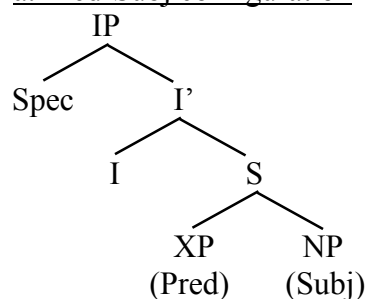
Kroeger (1993) proposes the principle in (145) for Tagalog clitic placement:

(145) Clitic positioning principle (Kroeger 1993:137)  
 Clitics occur immediately after the first (lexical or phrasal) daughter of the smallest maximal projection containing the head which governs them.

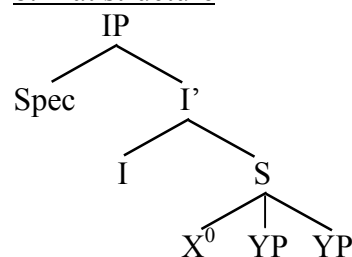
In order to understand how this principle functions it is first necessary to understand the phrase structure posited by Kroeger (1993:133) for Tagalog, shown in (146). Here, the S node does not indicate “Sentence” as in earlier transformational

work but rather “Small Clause”, and contains the predicate and its arguments in the spirit of Chung & McCloskey (1984). Two possibilities are available for the internal structure of S. In the configurational option, shown in (146)a, the predicate constitutes a phrase and is sister to the nominative subject. In the non-configurational option, the predicate head and its arguments are all direct daughters of S in a flat non-binary structure.

(146) a. Pred-Subj configuration



b. Flat structure



The flat structure is meant to explain facts both about phrasal syntax and clitic syntax. In phrasal syntax, non-configurationality allows a (full NP) nominative subject to intervene between a predicate head and its complement, as in (147), and in clitic syntax, non-configurationality predicts that clitics will follow the predicate head directly via (145).

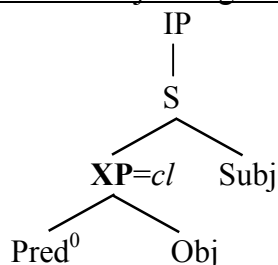
- (147) nag-bigay aŋ=paŋúlo naŋ=prémyo  
 AV.BEG-give NOM=president GEN=prize  
 ‘The president gave a prize.’

Given the possibility of non-configurationality, the configurational structure in (146)a would seem superfluous, as it can only yield a subset of the ordering possibilities allowed by the flat structure. Kroeger argues however on the basis of clitic positioning facts that it is necessary. Specifically, clitics are claimed to

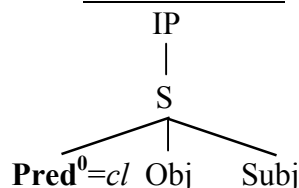
optionally follow the complement of a non-verbal predicate, a state of affairs which can only be explained by the phrasal constituency of [Pred<sup>0</sup> YP], assuming that clitics are positioned in accordance with (145).

Let us take a closer look at how (145) operates. The domain of clause-level pronominal arguments is IP (S being an exocentric category), and that of possessive pronouns, NP.<sup>64</sup> Discounting S because of its exocentric nature, clitics follow the left-branches of IP and stop at the first lexical XP or X<sup>0</sup> category, positioning themselves to the right of this. Assuming [Spec,IP] and I<sup>0</sup> to be empty, the clitics will be positioned after XP in the configurational structure in (148)a and after the predicate head in (148)b, as these are the first lexical constituents encountered on the left edge. The first daughter approach to 2P clitics has antecedents in the work of Klavans (1980 *et seq*) and Hale (1982) who attempted to substantiate an apparent relation between the presence of 2P clitics and free word order in their non-configurational analyses of Australian languages (see also Nevis 1986 and Taylor 1990).

(148) a. Pred-Subj configuration



b. Flat structure



<sup>64</sup> It is, in fact, unclear whether these domains are correctly predicted from the structural description “smallest maximal projection”. Kroeger (1993:137) takes the fact that non-subject argument clitics do not take the VP as their domain as evidence against the VP. However, this does not account for why the phrase indicated as XP in (146)a does not constitute a maximal domain for genitive clitics when the configurational option is selected (with non-verbal predicates, as verbal predicates are claimed to never employ the configurational structure). In other words, we should expect sentences such as (i) to be grammatical, in which a possessor clitic from a nominal predicate takes XP as its positioning domain.

- (i) \*?<sub>IP</sub>[Hindi=siya <sub>XP</sub>[kaibigan=ko]]  
 NEG=3S.NOM friend=1S.GEN  
 ‘He is not my friend.’



Kroeger claims an asymmetry between verbal and non-verbal predicates in that verbal predicates do not license the configurational structure.<sup>65</sup> Unfortunately, the empirical basis for post-phrasal positioning is very weak. The optionality of clitic placement with non-verbal predicates is based on a claim made by S&O (p.189-193) and further discussed in Sityar (1989): Verbal predicates, according to S&O, require clitics to follow the verb directly when it is clause initial but nominal and prepositional predicates allow delayed placement, more specifically, positioning after the predicate phrase. S&O (p.189-193) present the sentences in (149) as grammatical examples of post-phrasal positioning of clitics with non-verbal predicates. Kroeger (1993) interprets this to imply that non-verbal predicates can employ either the structure in (146)a or b. When these predicates are phrasal, the clitics surface following the oblique complement, as in (149). When they are flat, the clitics surface after the predicate head. Verb phrases, in contrast, have no phrasal option. They are always flat and, consequently, post-phrasal positioning is ungrammatical, as shown in (150), a putative minimal pair to (149)a.

- (149)a. Takot sa=kulog=**siya**                      b. Gáliñ sa=Maynìlà=**siya**  
 afraid OBL=thunder=3S.NOM                      from OBL=Manila=3S.NOM  
 ‘He is afraid of thunder.’                      ‘He is from Manila.’

- (150) na-tákot[=siya]                      sa=kulog[\*=siya]  
 PV.NVL.BEG-fear=3S.NOM OBL=thunder  
 ‘He was frightened by the thunder.’

I have found no evidence at all that the post-phrasal positioning in sentences like (149) is any better with non-verbal predicates than it is with verbal predicates.

---

<sup>65</sup> This stipulation seems unnecessary as Kroeger also assumes V to I movement which would render the verb the first lexical head within IP (and thus an obligatory clitic host) regardless of which structure was chosen.

Speakers find these structure to be, at the very least, highly awkward (see also Kroeger 1993:133, fn.13). These reactions were corroborated by the results of Google searches targeting exemplars of these structures. Post-head positioning of the 1S.NOM clitic in the same sentence yielded 16 hits while phrase final positioning was unattested. (Searching for the sentence in (i) verbatim only yielded citations of S&O’s original example in other linguistic literature.) Further results for different pronominal combinations are shown in Table 5.2. Strikingly, not a single example was found with clitics in post-phrasal position.

- (151)                      16 hits                      0 hits  
                                  takot[=ako]    sa=kulog[=ako]  
                                  afraid=1S.NOM OBL=thunder=1S.NOM

I thus take the putative categorial asymmetry originally claimed to exist by S&O (and repeated in Sityar 1989, Kroeger 1993, 1998, Carnie 1995) as mistaken. Post-phrasal positioning can only come about as a result of an extra layer of structure which correlates with a presuppositional interpretation (§5.5.1.3).

**Table 5.2.** Post-predicate head versus post-oblique phrase positioning of clitics

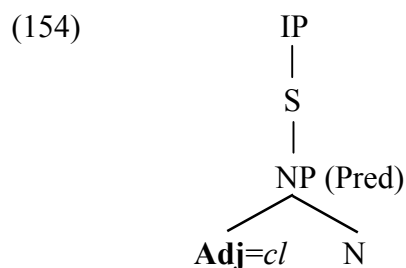
	OBL=1S	OBL=3S
1S.NOM	∅	149 hits      0 hits takot[=ako] sa=kanya[=ako]
2S.NOM	53 hit      0 hits takot[=ka] sa=(a)kin[=ka]	21 hits      0 hits takot[=ka] sa=kanya[=ka]
2P.NOM	7 hit      0 hits takot[=kayo] sa=(a)kin[=kayo]	21 hits      0 hits takot[=kayo] sa=kanya[=kayo]
3S.NOM	14 hits      0 hits takot[=siya] sa=(a)kin[=siya]	2 hits      0 hits takot[=siya] sa=kanya[=siya]
3P.NOM	70 hits      0 hits takot[=sila] sa=(a)kin[=sila]	9 hits      0 hits takot[=sila] sa=kanya[=sila]

More generally, taking the possibility of clitic intrusion to signal a flat syntactic structure makes it difficult to explain why other elements cannot also intrude in a similar fashion. Kroeger (1993:152-4) notes a problem in relation to intermediate X' constituents although only a tentative prosody based solution is suggested there. The difficulty is perhaps most apparent with complex nominal and adjectival predicates, which categorically disallow any intrusion into their domain from outside except for 2P clitics. The contrast between a clitic and full noun phrase subject with complex predicates is shown in (152) and (153).

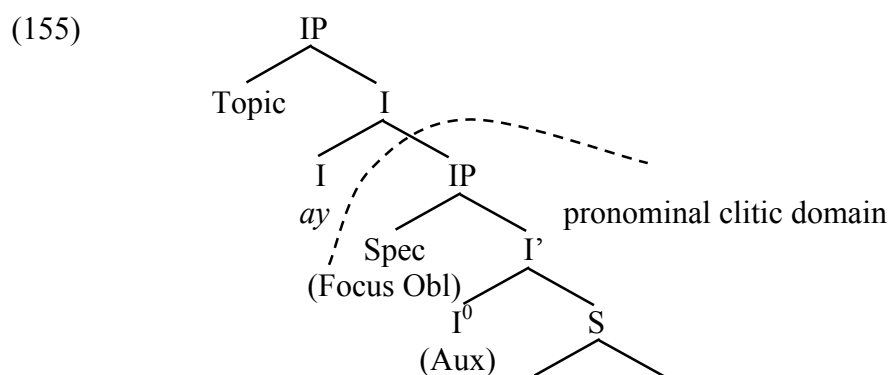
(152)a. importante=**ako**=η      táo      b. \*importante **añ=ámo**=η      táo  
 important=1S.NOM=LNK person      important      NOM=boss=LNK person  
 'I'm an important person.'

(153)a. masyado=**ka**=η      ma-ganda      b. \*masyado **añ=ámo**=η      ma-ganda  
 too=2S.NOM=LNK ADJ-beauty      too      NOM=boss=LNK ADJ-beauty  
 'You're too beautiful.'

Following Kroeger's analysis, the structure for (152)a would be as in (154). The smallest maximal projection of the lexical head in (154) is NP, yet, it is not this constituent which hosts clitics (in the unmarked case). For the clitic to not take the entire NP as its host, this constituent would have to be something other than an XP constituent but the data incontrovertibly suggests the existence of NP by all constituency diagnostics.



Kroeger's analysis of the left periphery and its relation to the pronominal clitic domain is shown in (155). Focused obliques appear in [Spec,IP] and are thus included within the same minimal IP as argument clitics but topics followed by the topic marker *ay* appear in the specifier of a higher adjoined IP and are thus outside the proper domain of clisis. I<sup>0</sup> may be filled by the auxiliary *huwag* PROHIBITIVE or by the verb itself. The relative positions of topic and oblique focused constituents is supported by prosodic parsing in that, unlike the situation with topics, it is impossible to pause after a focused oblique (S&O:493,485,496).



The predictions of the structure in conjunction with the clitic positioning principle in (145) is that focused oblique phrases in [Spec,IP] will act as unitary hosts for clitics and that only one of these constituents should appear. Kroeger explicitly likens the [Spec,IP] position in Tagalog to the Germanic topic position. It is claimed to be similar in that it is unique but differs in being only optionally filled and carrying a strong focus reading. In both Germanic and Tagalog, these constituents play host to a 2P element, the verb in Germanic and argument clitics in Tagalog.

Note also that given Kroeger's clitic positioning principle in (145), we predict that multiple adjunction to IP should result in clitics being placed after the lowest adjunct, as the lowest adjunct would represent the smallest maximal projection of the

predicate head. But when both IntP and FocP are occupied by oblique phrases, we find that the higher of the two must host the clitic, as seen in (156).

- (156) saan=**ka**            sa=Manílà nakatira?  
 where=2S.NOM OBL=M.    live  
 ‘Where in Manila do you live?’<sup>66</sup>

Finally, Kroeger claims that impenetrability is due to the fact that the fronted oblique is housed in [Spec,IP] and therefore must be treated as a singular constituent for the purposes of clitic placement, as it is the first syntactic daughter within the proper domain. Kroeger offers independent evidence that all material in the focused oblique position is a singular constituent in [Spec,IP] from interactions with negation. This is seen in (157) (Kroeger’s 1993:130). The negative polarity item *káhit ninuman* ‘whoever GEN’ is interpreted as ‘no one’ when it falls under the scope of negation and ‘any one’ when it does not. In (157)a, clause initial negation licenses the negative polarity reading and the subject is interpreted as ‘no one at all’. In (157)b we find the same interpretation except that the focused oblique lies outside the scope of negation. In (157)c, the pivotal example, the negative polarity reading of *káhit ninuman* is not licensed when negation is in [Spec,IP], rather negation has narrow scope over the focused adjunct. The point is further driven home by (157)d where we see that double negation, which is otherwise ungrammatical in Tagalog, is allowed when one of the negation elements is in [Spec,IP].

- (157) a. Hindi=siya    k<in>á~ka-úsap            naḡ=káhit ninuman sa=opisína  
 NEG=3S.NOM <BEG>~INCM-co-speak GEN=even whoever OBL=office  
 ‘No one at all talks to him at the office.’
- b. [Sa=opisina]=siya    hindi k<in>á~ka-úsap            naḡ=káhit ninuman  
 OBL=office=3S.NOM NEG <BEG>~INCM-co-speak GEN=even whoever  
 ‘At the office no one at all talks to him.’

<sup>66</sup> From: [www.suncatcherph.blogspot.com/2005\\_08\\_01\\_archive.html](http://www.suncatcherph.blogspot.com/2005_08_01_archive.html)

- c. [Hindi sa=opisina] siya k<in>á~ka-úsap naŋ=káhit ninuman  
 NEG OBL=office 3S.NOM <BEG>~INCM-co-speak GEN=even whoever

(péro sa=eskwela)

but OBL=school

‘It’s not at the office that everyone talks to him (but at school).

Not: \*At the office no one at all talks to him.

- d. [Hindi sa=opisina] siya hindi k<in>á~ka-úsap naŋ=káhit ninuman  
 NEG OBL=office 3S.NOM NEG <BEG>~INCM-co-speak GEN=even whoever  
 ‘It’s not at the office that no one at all talks to him.’

However, the fact that negation can form a constituent with a focused oblique as in (157)c and d, does not imply these elements *must* form a constituent in this configuration. Indeed, when negation precedes a focused oblique the clitic may also attach to negation, as shown in (158). This is problematic because if the [Spec,IP] position is unique, as claimed, and clitics are bound within their minimal IP, then the post-negation position below should be impossible.

- (158) ...hindi=ka sa=ákin maka~ka-hinì  
 NEG=2S.NOM OBL=1S AV.NVL~INCM-request  
 ‘...you won’t be able to request from me.’<sup>67</sup>

The same holds true for cases of an adjunct interrogative preceding a focus fronted oblique phrase, as in (159). Just as in (158), the clitic must attach to the interrogative.<sup>68</sup>

- (159) Kailan[=ka] sa=akin[?\*=ka] Ø-dá~dálaw?  
 when=2S.NOM OBL=1S=2S.NOM AV-INCM~visit  
 ‘When will you visit me?’

<sup>67</sup> From: [http://www.lawphil.net/judjuris/juri1995/sep1995/gr\\_93833\\_1995.html](http://www.lawphil.net/judjuris/juri1995/sep1995/gr_93833_1995.html)

<sup>68</sup> Some speakers reject an interrogative cooccurring with a focus fronted oblique phrase. Note that this environment is similar to the Bulgarian example cited by Bošković (2002) (see also Rudin 1988) in which clitics must follow the first wh- element when multiple wh- elements are present.

- (i) CP[Koliko im XP[ko TP[daje ...  
 How-much them who gives  
 ‘Who gives them how much?’

The difficult data with multiple fronted constituents is accounted for here without further ado by the phrase structure proposed in (1) above in conjunction with the Clitic Visibility Constraint. It has been furthermore shown here that the categorial asymmetries claimed by S&O, Kroeger and Sityar are seen to result from presuppositionality effects rather than different syntactic configurations for different lexical categories.

### 5.9.2 *Billings & Konopasky: verb adjacency*

Billings & Konopasky (2003:32) claim that, “An overview of the data in Schachter & Otnes (1972) and Kroeger (1993) reveals that the clitics are also invariably verb-adjacent (assuming that negation is part of the cluster of clitics, as in Bulgarian).” They compare the Tagalog facts to those of Bulgarian in (160) and (161) (Billings & Konopasky 2003:21). Bulgarian pronominal clitics are verb-adjacent and show Tobler-Mussafia effects such that they are proclitic unless the verb is clause initial in which case they are enclitic. Based solely on the data in (161), Tagalog could instantiate the same pattern.

(160) a. Az ti=gi=dadox                      b. Dadox=ti=gi                      *Bulgarian*  
           1s.NOM 2s.IO=3p.DO=gave.1s        gave.1s=2s.IO=3p.DO  
           ‘It’s me that gave them to you’      ‘I gave them to you.’

(161)a. Kahápon=ka=ba=nila                      na-kítà  
           yesterday=2s.NOM=QM=3p.GEN PV.NVL=see  
           ‘Was it yesterday that they saw you?’

b. Na-kítà=ka=ba=nila                      kahápon  
           PV.NVL=see=2s.NOM=QM=3p.GEN yesterday  
           ‘Did they see you (yesterday)?’

Unfortunately, Billings & Konopasky do not explore more complex examples involving multiple potential interveners and thus do not provide the necessary evidence for verb-adjacency over second-position. This is especially important given that Tagalog is a predicate initial language and therefore the verb is often proximate to the left edge of the clause, making the two possibilities difficult to tease apart. In fact, verb-adjacency, or predicate adjacency may be one factor in determining the choice of optional orders, however it is far from a categorial constraint on clitic positioning. With multiple pre-predicate elements, we often find both verb-adjacent orders, as in (162), and second position orders, as in (163).

(162) bigla=η            hindi=**ako**            k<in>á~ka-úsap  
 sudden=LNK    NEG=1S.NOM=LNK    <BEG>co~INCM-converse  
 ‘all of a sudden, I wasn’t being spoken with.’<sup>69</sup>

(163) biglà            **ako**=η            hindi    naka-hiņa  
 sudden 1S.NOM=LNK    NEG    AV.NVL.BEG-breathe  
 ‘all of a sudden, I wasn’t able to breathe.’<sup>70</sup>

A focus fronted oblique can also intervene between an interrogative and the verb as show in (159) above and (164) below.

(164) saan=**ka**            sa=Manilà    nakatira?  
 where=2S.NOM    OBL=M.    live  
 ‘Where in Manila do you live?’<sup>71</sup>

Verb-adjacency could conceivably be argued to play an emergent role in the orderings shown in the minimal pair (165) and (166). Although both orders are accepted by speakers, the verb-adjacent order in (166) has a much higher rate of attestation in natural written text: the string “*madalas ako hindi*” (often *cl* NEG) with

<sup>69</sup> From: [www.mia-rinascita.livejournal.com/tag/broad](http://www.mia-rinascita.livejournal.com/tag/broad)

<sup>70</sup> From: [http://patricia09.multiply.com/journal/item/89/Bigla\\_akong\\_hindi\\_nakahinga](http://patricia09.multiply.com/journal/item/89/Bigla_akong_hindi_nakahinga).

<sup>71</sup> From: [www.suncatcherph.blogspot.com/2005\\_08\\_01\\_archive.html](http://www.suncatcherph.blogspot.com/2005_08_01_archive.html)



second position ordering returned only one legitimate hit on Google while the verb-adjacent order “*madalas hindi ako*” (often NEG *cl*) yielded 70 hits (although not all legitimate instantiations of the target structure).

(165) *ma-dalas=ako hindi maka-túlog.*  
 ADJ-often=1S.NOM NEG AV.NVL-sleep  
 ‘Often, I don’t sleep.’<sup>72</sup>

(166) *ma-dalas hindi=ako maka-túlog*  
 ADJ-often NEG=1S.NOM AV.NVL-sleep  
 ‘Often, I don’t sleep.’<sup>73</sup>

On the other hand, for the string in which negation precedes an adverb (with the corresponding scope NEG>ADV), as in (167), the results overwhelmingly favored positioning in the 2P pattern rather than the verb-adjacent one. Speakers also feel that verb adjacent ordering in such sentences is awkward to ungrammatical.

(167) *Hindi=ako ma-dalas naki~kí-pag-úsap.*  
 NEG=1S.NOM ADJ-often AV.BEG.SOCL~INCM-TR-converse  
 ‘I don’t often converse (with people)’<sup>74</sup>

One possible reason for the strong tendency towards verb-adjacency in ADV + NEG orders of the type (166) is that frequency adverbs tend to be topicalized when they have wide scope over other operators. This is made more transparent by the orthography of several attestations which employ a comma between the frequency adverb and negation, as the one in (168). As discussed above in §5.3.1, topics can never host argument clitics and thus the post-negation position is expected if a preceding adverb is topicalized. (Note that, despite the apparent delayed clitic

<sup>72</sup> From: [www.orengyouglad.tabulas.com/2006/06/04/@1215044/](http://www.orengyouglad.tabulas.com/2006/06/04/@1215044/)

<sup>73</sup> From: [http://bonnarenz.blogs.friendster.com/my\\_blog/2007/09/daily\\_routine.html](http://bonnarenz.blogs.friendster.com/my_blog/2007/09/daily_routine.html)

<sup>74</sup> From: [www.profiles.friendster.com/13288347](http://www.profiles.friendster.com/13288347)

placement in (168), it still does not instantiate verb-adjacent clisis because of the intervening focused oblique.) On the other hand, negation can never be topicalized which explains why negation must always host clitics in structures such as (167).<sup>75</sup>

(168) ma-dalas, hindi=**ako** sa=*camera* naka-tiñin  
 ADJ-often NEG=1S.NOM OBL=*camera* AV.NVL-look  
 ‘Often, I’m not looking at the camera.’<sup>76</sup>

In order to account for some of the exceptions, Billings & Konopasky (2003) make an unorthodox assumption concerning negation in Tagalog, namely, that it should be considered as part of the clitic cluster, thus allowing the string *cl* NEG VERB to satisfy verb-adjacency. This assumption appears to have been made solely on the analogy of Bulgarian and certain other Slavic languages in which negation has been argued to form a constituent with the verb. In Tagalog, however, there is nothing at all to suggest that negation is clitic-like. Negation is a free-standing element which can independently form a full-utterance and which never shows 2P effects.

### 5.9.3 *Balkan DP-internal clitics in Distributed Morphology*

Several of the properties related concerning clitic placement in the Tagalog DP are reflected very closely by their counterparts in the Balkan sprachbund. Standard Modern Greek preserves historically more general 2P effects in the nominal domain unlike in the more innovative clausal domain which only attests verb adjacent clitics. As in the Tagalog (53) above, Greek also allows both post-modifier and post-nominal positioning of the genitive clitic when the nominal head is preceded by a modifier, as shown in (169), but only post-phrasal positioning of full NPs, as shown in (170).<sup>77</sup>

<sup>75</sup> Recall from the previous chapter that it was precisely when verbs followed negation that pronominal enclisis to the verb was impossible in Old Catalan, a fact which was also attributed to the same reason.

<sup>76</sup> From: [www.twenty2two.multiply.com/photos/album/26](http://www.twenty2two.multiply.com/photos/album/26)

<sup>77</sup> All Greek data is taken from Alexiadou & Stavrou (2000) who, by way of arguing for an analysis of Greek possessor clitics, also discuss certain subtle differences in interpretation between the clitic

(169) to paljo[=mu] aftokinito[=mu] *Modern Greek*  
 the old=1S.GEN car  
 ‘my old car’ (Alexiadou & Stavrou 2000)

(170) to kenurjio [\*tu pedhju] podhilato [tu pedhju] *Modern Greek*  
 the new the-GEN kid-GEN bike  
 ‘The kid’s new bike’ (ibid.)

Furthermore, when the nominal precedes the modifier, as in (171), putting the clitic in post-modifier position is ungrammatical, just as seen earlier in Tagalog (56).

(171) to vivlio[=mu] to kenurjo[\*=mu] *Modern Greek*  
 the book=1S.GEN the new  
 ‘My new book/the new book of mine’ (ibid.)

Finally, placement of a genitive clitic within a complex modifier phrase, as with *kapos ipervolikos* ‘somewhat excessive’ in (172) and *iperifanos jia ta pedhja* ‘proud of his children’ in (173), is illicit.

(172) o kapos[\*=tis] ipervolikos[=tis] enthusiasmos *Modern Greek*  
 the somewhat=3S.F.GEN excessive enthusiasm  
 ‘her somewhat excessive enthusiasm’ (ibid.)

(173) o iperifanos[\*=tis] jia ta pedhja=tu pateras[=tis] *Modern Greek*  
 the proud=hers for the children=his father  
 ‘her father who is proud of his children’ (ibid.)

Embick & Noyer (1999, 2001) discuss similar facts from Bulgarian (see also Halpern 1995), shown in (174). Here, the definite article *ta* appears, generally speaking, after the first available host in DP. In (174)a, this is the noun itself, in

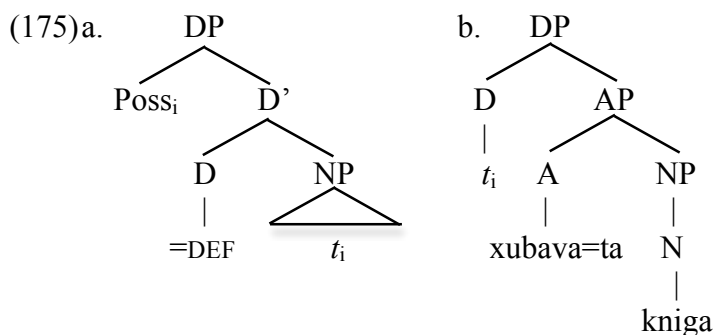
---

positions in (169). The translations have been simplified for our purposes here (and also in accordance with the opinions of several native speakers who could not affirm the claimed distinctions) and the glossing has been adapted to present conventions.

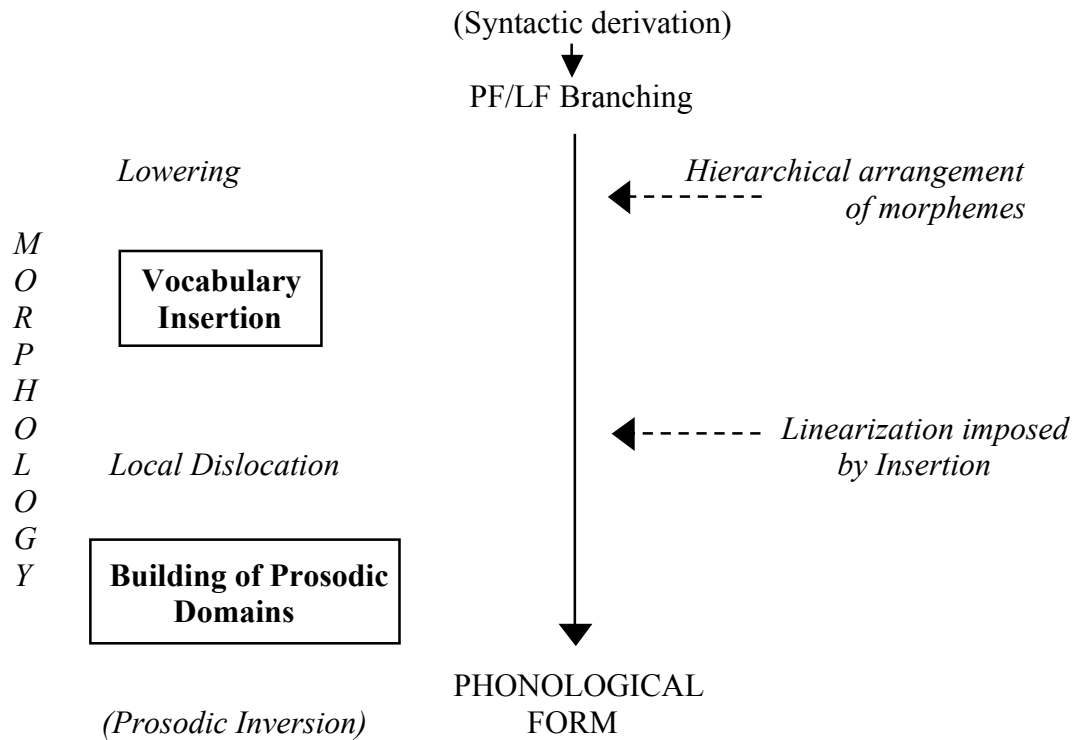
(174)b, a prenominal adjective and in (174)c it is a possessive pronoun which precedes the adjective.

- (174)a. kniga=**ta**  
 book=DEF  
 ‘the book’
- b. xubava=**ta** kniga  
 nice=DEF book  
 ‘the nice book’
- c. moja=**ta** xubava kniga  
 my=DEF nice book  
 ‘My nice book’

Their analysis of the 2P effects here is reliant on “Lowering”, an operation conceived of as taking place in the morphology. They claim that the possessive raises from a lower position into [Spec,DP], from which it can host the definite clitic while the clitic remains in its base position, as shown in (175)a. When the adjective intervenes, the definite article lowers onto the head, as illustrated in (175)b.



Embick & Noyer do not discuss the trigger for Lowering nor why it is only triggered when no preceding material is available. It must thus be assumed that the trigger is phonological, i.e., due to the lack of an appropriate host for enclisis. But this is problematic for their model, shown in Figure 1 (repeated from §2.3.4), as Lowering precedes Vocabulary Insertion and should have no access to phonological information.



**Figure 5.1.** DM architecture (Embick & Noyer 1999:273)

The process of Local Dislocation takes place after Vocabulary Insertion, but unlike, Lowering, it is restricted to adjacent material and can be blocked by intervening elements. Thus, while Local Dislocation could make the required reference to phonology it cannot be responsible for the Bulgarian positioning facts, since, just as in Tagalog and Greek, clitics must be able to “skip” adverbial modifiers, as shown in (176).

- (176) mnog[\*=ət] star[=ət] teatər  
 very=DEF old theater  
 ‘the very old theater’

Skipping of the adverb in (176) represents non-local reordering and can thus only be an instantiation of Lowering, an operation which targets syntactic heads:

“The Definite element Lowers across intervening material to the immediately dominated head, subject to certain morphophonological requirements. The fact that DEF does not appear on the adverb is then purely structural; when DEF undergoes Merger, it targets the head of its complement, stated in terms of syntactic headedness. The adverb, being an adjunct, cannot be the target of this operation.”

(Embick & Noyer 1999:276)

This analysis of the impenetrability facts crucially depends on a particular structural interpretation of adjectival and adverbial modification, namely, one in which adjectives are heads selecting for NP complements (Abney 1987) and adverbs are adjuncts to the adjectival phrase. Problematically, there is nothing head-like about the relationship of adjectives to the noun phrases they modify (Hankamer & Mikkelsen 2005, Dost & Gribanova 2006, Kramer to appear). It also remains unclear why preceding material (e.g., possessor pronouns) should bleed what is conceived of as the definite article’s morphological subcategorization.<sup>78</sup> In other words, if Lowering satisfies morphological requirements of the definiteness morpheme we expect it to take place whether or not [Spec,DP] is occupied by a possessor.

The pattern found with Macedonian is argued by Embick & Noyer (1999) to form a minimal pair with Bulgarian. In Macedonian, the definite article also attaches to the first element within the NP, as shown in (177). Macedonian is claimed to differ, however, in not allowing the definite clitic at all when a prenominal adjective is modified by an adverbial, as shown in (178). Unlike Bulgarian, not only can the definite article not attach to the adverbial itself, as in (178)a, it cannot skip the adverbial to attach to the following adjective, as in (178)b. The solution, according to

---

<sup>78</sup> This is made explicit by their “Local Dislocation Hypothesis”, which states that, “If a movement operation is vocabulary sensitive, it involves only string-adjacent items.” (Embick & Noyer 1999:274). They elaborate, “Where a clitic demands a host having a particular identity, such as inflectional class, morphological category, or phonological weight, then in our terms the operation is vocabulary-sensitive and the clitic and the host must be string-adjacent prior to Merger.”

Sadock (1991, via Victor Friedman p.c.), is to use a demonstrative in place of the definite article, as in (178)c.

- (177) a. *čovек=ot*                      b. *dobr=iot čovek*                      c. *dobr=iot mal čovek*  
           man=DEF                              good=DEF man                              good=DEF little man  
           ‘the man’                              ‘the good man’                              ‘the good little man’
- (178)a. *\*mnogu=ot/ta/to/te* *golem čovek*                      b. *\*mnogu golem=iot čovek*  
           very=DEF                              big    man                                      very    big=DEF    man
- c. *onoj mnogu golem čovek*  
           that very    big    man  
           ‘the very big man’

Embick & Noyer (1999) treat the difference between Bulgarian and Macedonian as one between Lowering (for Bulgarian), which ignores intervening material and takes place before Vocabulary Insertion, and Local Dislocation (for Macedonian), which is sensitive to intervening material and constituency.

Unfortunately, the empirical data is not at all clear. In a later grammar sketch by Victor Friedman, to whom the original observation was attributed, we find no mention of the putative rule. Rather, he describes the situation much like Bulgarian, “The article attaches to the end of the first nominal in the noun phrase, that is nouns, adjectives, pronouns, numerals, but not adverbs.” (Friedman n.d.:17). More to the point, however, in demonstrating the placement of the definite article in complex NPs, he offers examples of the putatively ungrammatical kind, shown in (179) and (180).<sup>79</sup>

- (179) *ne mnogu postari=te deca*  
           not much    older=DEF children:PL  
           ‘the not much older children’ = ‘the children that are not much older’
- (180) *edna od mnogu=te naši zadači*  
           one from many=DEF our:PL problems:PL  
           ‘one of our many problems’

---

<sup>79</sup> Note that the use of *mnogu* in (180) is not adverbial, as it is in (179), but rather adjectival.





adjectives, they are invisible to the external nominal and thus ruled out.<sup>81</sup> This pattern, in fact, seems extremely widespread. Kramer (to appear) discusses identical impenetrability facts for the 2P definite determiner in Amharic.<sup>82</sup> Although, the case of Amharic is complicated slightly by multiple occurrence of the definite marker (treated by Kramer as definite agreement), we can observe the similarities in the simple cases shown in (183). The definite marker, /u/ post-vocally and /w/ post-consonantly, surfaces after a noun when nothing precedes within the NP, as in (183)a, but after a modifier when it precedes the head noun. This can be seen with a simple adjective in (183)b and a relative in (183)c.

---

<sup>81</sup> There are, however, some Bulgarian examples which appear problematic for the CVC but which cannot be fully discussed here. Petcova & Alahverdzhieva (2005) cite examples (i)-(iii) as cases where the definite article can appear sandwiched within a modificational phrase. It is possible that definiteness, unlike thematic role assignment, need not be a relationship between the clitic and the lexical head but can also target other elements within the NP. If this is correct it must be the case that *poluchena=ta*, *verni=jat* and *kupeni=te* are all possible independently.

- (i) poluchena=**ta** s=maka stipendij  
received=DEF with=pain scholarship  
'The received with pain scholarship'
- (ii) verni=**jat** na=demokratichni idei prezident  
faithful=DEF to=democratic ideas president  
'the president (who is) faithful to democratic ideas'
- (iii) kupeni=**te** vchera knigi  
bought=DEF yesterday books  
'the books (which were) bought yesterday'

Note though that placement here is perhaps more sensitive to hierarchical structure. This can be seen by the fact that attachment to embedded adjuncts is impossible regardless of their surface position, as shown by Tomić (1996b:521):

- (iv) sās māka polučena=**ta** stipendija  
with pain obtained=DEF scholarship  
'painfully obtained scholarship'
- (v) \*sās māka=**ta** polučena stipendija  
with pain=DEF obtained scholarship  
'painfully obtained scholarship'

<sup>82</sup> Kramer's analysis of Amharic, which I have come across too late to discuss fully here, bears some similarities to that argued for here. Specifically, she argues that the definiteness clitic must be in the spell-out domain of DP for reasons of the Phase Impenetrability Condition.

- (183) a. *bet*=**u**                      b. *tillik'*=**u** *bet*                      *Amharic*  
house=DEF                      BIG=DEF house  
‘the house’                      ‘the big house’
- c. *yä-särräk'-ä*=**w**      *ličč*  
C-steal.PF-3MS=DEF child  
‘the child who stole’

However, when the modifier is complex, as in the examples in (184), it constitutes an impenetrable constituent for the purposes of clitic placement.

- (184) a. [*tinantina yä-mät't'-a*]-**w**      *tämari*                      *Amharic*  
yesterday C-come.PF-3MS-DEF student  
‘the student who came yesterday’
- b. [*iččig bät'am tillik'*]=**u** *bet*      c. [*lä-mist-u tammaññ*]=**u** *gäs'äbahriy*  
really very big=DEF house      to-wife-his faithful=DEF character  
‘the really big house’                      ‘the faithful-to-his-wife character’

The surface generalization is thus identical to that of Tagalog and the Baltic examples and can be handled equally by the Clitic Visibility Condition. The CVC avoids having to posit the unattractive analysis of adjectives as heads taking NP complements (see Hankamer & Mikkelsen 2005 for arguments against this) and need not stipulate non-attachment to adverbs as a lexical class (as in Dost & Gribanova 2006). This latter move appears unmotivated considering that adjectives and adverbs do not appear to be well defined lexical classes in the languages under consideration, as seen by the two uses of *mnogu* ‘many’ (179) and (180). It is only the syntax which can determine that the former is adverbial and the latter is adjectival. Similarly, it is only the syntactic embeddedness of adverbs which makes them bad hosts in all of the above languages.

## **5.10 Conclusion**

This chapter has offered a detailed look at various aspects of Tagalog syntax and the interactions of clitic placement in various constructions. We identified the range of clitic placement as the spell out domain of the containing phase and argued for the Clitic Visibility Condition which is based on the notion of Linear Containment. This was shown to handle the facts better than previous theories which make direct reference to syntactic structure, lexical category, or pure prosody. Along the way, we have elucidated several aspects of Tagalog syntax relating to the left periphery of the sentence and DP internal elements.

## CHAPTER 6: CONCLUSION

### 6.1 Summary and prospects

I have argued here on the basis of a detailed examination of Tagalog as well as a review of the typological literature that clitics come in two flavors: those which are bona-fide syntactic elements generated as part of the phrase structure and those which are the spell out of feature bundles adjoined to phrase structure. This distinction allows us to make far better sense of cross-linguistic typology. When clitics are under phonological pressure to avoid initial position in a prosodic domain, “syntactic” clitics will attach to a preceding constituent while “featural” clitics will overwhelmingly tend to take second position. Prosodic movement can thus be limited in its scope to elements which do not possess a terminal in phrase structure. Bona-fide syntactic heads on the other hand can only be manipulated by syntactic movement, explaining their absence in second position.

While the introduction of syntactic elements into the derivation has been understood here to take place via a standard Merge operation, it has been argued that feature bundles may be merged without a terminal node, adjoined directly to phrase edges. This not only goes a long way in explaining their prosodic and positioning behavior but was also shown to account for the sensitivity of clitics to a variety of morphological phenomena (idiosyncratic linearization, cooccurrence constraints) which do not effect free pronouns.

Regarding impenetrability phenomena, we found that there exists a PF locality condition between pronominal clitics and the predicate phrase from which they are assigned a theta role. This relationship resembles the ECP but crucially pertains to clitics which demonstrably lack a syntactic position. The solution here made use of

a purely linear notion: Linear Containment, to capture restrictions on elements which are linearized between nodes.

The approach taken in this dissertation has been completely representational in nature on the assumption that a derivational approach employing the same richness of representations would unnecessarily complicate matters. More importantly, it has been taken as a methodological tack in order to capture as many transparent surface generalizations as possible. Nonetheless, this is clearly not the only approach. One area in particular which may be profitably reanalyzed in more derivational terms is the Phase Correspondence Principle (§4.2.1). Following Fox & Pesetsky (2004) and other works in the same vein (cf. Kramer to appear), the cyclic linearization of phases may offer a unified explanation for domain phenomena (the inability of TP clitics to lean on material in  $C^0$ ) as well as impenetrability phenomena (if focused obliques can be argued to be phases). There are several hurdles which such an analysis must face, but the pay off is potentially great. Hopefully, in addition to furthering our understanding of Tagalog and 2P phenomena, the generalizations discovered here will be able to open up new vistas for exploring cyclicity at the phonology-syntax interface.

## REFERENCES

- Abdulla, Jamal, and Ernest McCarus. 1967. *Kurdish Basic Course*. Ann Arbor: University of Michigan Press.
- Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect*. Ph.D. dissertation, MIT.
- Adelaar, K. Alexander and Nikolaus Himmelmann. 2005. *The Austronesian languages of Asia and Madagascar*. Routledge language family series 7.
- Adger, David. 1997. VSO order and weak pronouns in Goidelic Celtic. *Canadian Journal of Linguistics* 42:9–29.
- Adriani, Nicolaus, and Samuel Esser. 1939. *Koelawische taalstudien. Vol. 1. Overzicht der spraakkunst, gesprekken en verhalen met vertaling*. Bandoeng: Nix.
- Agbayani, Brian and Chris Golston. To appear. Second-position is first-position: Wackernagel's Law and the role of clausal conjunction. *Diachronica*.
- Aissen, Judith. 1992. Topic and Focus in Mayan. *Language* 68:43-80.
- Alan Prince & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report 2.
- Aldridge, Edith. 2002. Nominalization and WH-movement in Seediq and Tagalog. *Language And Linguistics* 3.2:393-426.
- Aldridge, Edith. 2004. *Ergativity and Word Order in Austronesian Languages*. Ph.D. dissertation, Cornell University.
- Alexiadou, Artemis & Melita Stavrou. 2000. Adjective-Clitic Combinations in the Greek DP. In Birgit Gerlach and Jan Grijzenhout (eds.) *Clitics in Phonology, Morphology and Syntax*, 63-84. Amsterdam: John Benjamins.
- Anderson, Stephen R. 1972. On nasalization in Sundanese. *Linguistic Inquiry* 3:253-268.
- Anderson, Stephen R. 1982. Where's Morphology? *Linguistic Inquiry* 13: 571-612.
- Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

- Anderson, Stephen R. 1993. Wackernagel's Revenge: Clitics, Morphology, and the syntax of Second Position. *Language* 69:68-98.
- Anderson, Stephen R. 1996. How to Put Your Clitics in Their Place. *The Linguistic Review* 13:165-191
- Anderson, Stephen R. 2005. *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.
- Arka, Wayan and J. Kosmas. 2005. Passive without passive morphology? Evidence from Manggarai. In I Wayan Arka and Malcolm D. Ross (eds), *The many faces of Austronesian voice systems: some new empirical studies*, 87-117. Canberra: Pacific Linguistics.
- Ba'dulu, Abd. Muis. 1990. Frase verbal bahasa Mandar Majene. In Husan Abas and T. David Andersen (eds.) *Bahasa-bahasa daerah Sulawesi dalam konteks bahasa nasional*, 69-84. Ujung Pandang: Universitas Hasanuddin and Summer Institute of Linguistics.
- Baker, Mark. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373-415.
- Baker, Mark. 1988. *Incorporation*. Chicago: Chicago University Press
- Baker, Mark. 1995. On the absence of certain quantifiers in Mohawk. In Emond Bach, Eloise Jelinek, Barbara Partee (eds.), *Quantification in natural languages*, 21-58. Dordrecht: Kluwer.
- Baker, Mark. 2005. *Lexical Categories: Verbs, Nouns and Adjectives*. Cambridge University Press.
- Baker, Mark. 2008. *The Syntax of Agreement and Concord*. Cambridge University Press.
- Basri, Hasan. 1999. *The syntactic and phonological reflections of morphological structure in Selayarese*. Ph.D. dissertation, SUNY Stony Brook.
- Basri, Hasan, Ellen Broselow, Dan Finer, and Elisabeth Selkirk. 1998. Morphologically Grounded Affix Classes: the Case of Makassarese. Ms. State University of New York at Stony Brook, NY and University of Massachusetts at Amherst, MA
- Beckman, Jill. 1998. *Positional Faithfulness*. Ph.D. dissertation, University of Massachusetts at Amherst.

- Bickel, Balthazar and Johana Nichold. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*. Cambridge: Cambridge University Press.
- Billings, L.A., 2002. Phrasal clitics. *Journal of Slavic Linguistics* 10:53-104.
- Billings, Loren A. 2005. Ordering clitics and postverbal R-expression in Tagalog: A unified analysis? In Andrew Carnie, Heidi Harley, and Sheila Ann Dooley (eds.). *Verb First: On the Syntax of Verb-initial Languages*, 303-39. (Linguistik Aktuell, 73.) Amsterdam: Benjamins.
- Billings, Loren and Abigail Konopasky. 2002. Morphology's role in ordering verb-adjacent clitics. In Line Mikkelsen and Christopher Potts (eds.) *The proceedings of the Twenty-first West Coast Conference on Formal Linguistics*, 29-42. Somerville, MA: Cascadilla Press.
- Billings, Loren and Abigail Konopasky. 2003. Reassessing the role of syntax inside the morphological word: verb-adjacent clitics in Tagalog and Bulgarian. In Anastasia Riehl, and Thess Savella (eds.), *Proceedings of AFLA 9*, Ithaca, NY, 2002. Cornell Working Papers in Linguistics 19.
- Billings, Loren, and Daniel Kaufman. 2004. Towards a typology of Austronesian pronominal clisis. *ZAS Papers in Linguistics* 34, 15-29. (Proceedings of AFLA 11.)
- Billings, Loren, Daniel Kaufman and Adam Werle 2006. Rightward-subordinating enclitics in Wakashan and Austronesian languages. Talk presented at the 2<sup>nd</sup> Wakashan Linguistics Conference. University of Victoria.
- Bloomfield, Leonard. 1917. *Tagalog texts with grammatical analysis*. University of Illinois studies in language and literature, 3.2-4. Urbana, Illinois.
- Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73-129.
- Blust, Robert. 2005. The linguistic macrohistory of the Philippines: Some speculations. In Hsiu-chuan Liao and Carl R. Galvez Rubino (eds.) *Current issues in Philippine linguistics and anthropology*, 31-68. Manila: The Linguistic Society of the Philippines and SIL Philippines.
- Börjars, Kersti. 1998. *Feature distribution in Swedish noun phrases*. Oxford: Blackwell.
- Bošković, Željko. 1995. Participle movement and second position cliticization in Serbo-Croatian. *Lingua* 96:245-266.



- Bošković, Željko. 2001. *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Elsevier.
- Bowern, Claire. 2004. *Bardi Verb Morphology from a Historical Perspective*. Ph.D. dissertation, Harvard University.
- Bowers, John. 1988. Extended X-Bar Theory, the ECP, and the Left Branch Condition. In *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL) 7*, 47-62, Stanford Linguistics Association, Stanford University.
- Bradley, Travis G. 2001. A Typology of Rhotic Duration Contrast and Neutralization. *Proceedings of the North East Linguistic Society 31*. Available on the Rutgers Optimality Archive: <http://roa.rutgers.edu/files/436-0601/436-0601-BRADLEY-0-0.PDF>
- Brainard, Sherri and Ena Vander Molen. 2005. Word order inverse in Obo Manobo. In Hsiu-chuan Liao and Carl R. Galvez (eds.), *Current issues in Philippine linguistics and anthropology parangal kay Lawrence A. Reid*, 364-418. Manila, Philippines: Linguistic Society of the Philippines and SIL Philippines.
- Bresnan, Joan & Sam Mchombo 1987. Topic, pronoun and agreement in Chichewa. *Language* 63:741-82
- Bresnan, Joan and L. Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21:147-185.
- Bresnan, Joan and Sam Mchombo. 1986. Grammatical and anaphoric agreement. *Papers from the Parasession on Pragmatics and Grammatical Theory*. Chicago Linguistic Society 22:2, 278-97.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*, Oxford: Blackwell Publishers Ltd
- Browne, Wayles. 1974. On the Problem of Enclitic Placement in Serbo-Croatian. In R. D. Brecht and Catherine Chvany (eds.), *Slavic Transformational Syntax*, 36-52. Michigan Slavic Materials 10, Ann Arbor.
- Cardinaletti Anna and Michal Starke. 1994/1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.), *Clitics in the Languages of Europe*, EALT/EUROTYP 20-5, 145-233. Mouton, Berlin-New York.
- Cardinaletti, A. & Roberts, I. 1991, Clause structure and X-second. Ms., Università di Venezia and Université de Genève.
- Carnie, Andrew. 2000. On the notions XP and X<sup>0</sup>. *Syntax* 3.2:59-106.

- Ćavar, Damir. 1996. On Clitics in Croatian: Syntax or Phonology? *ZAS-Working Papers in Linguistics* 6. pp. 51-65.
- Ćavar, Damir. 1999. *Aspects of the Syntax-Phonology Interface*. Ph.D. dissertation, University of Potsdam.
- Cheng, Lisa L.-S. 1991. *On the typology of wh-questions*. Ph.D. dissertation, MIT. Published 1997, Garland, New York
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.
- Chomsky, Noam. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds.), *Readings in English Transformational Grammar*, Blaisdell, Waltham, MA.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in comparative grammar*, 417-454. Cambridge: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In: M. Kenstowicz (ed.), *Ken Hale: A life in language*, 1-52. Cambridge, Mass.: MIT Press.
- Chung, Sandra, and James McCloskey. 1987. Government, barriers and small clauses in Modern Irish. *Linguistic Inquiry* 18:173–237.
- Chung, Sandra. 1998. *The Design of Agreement: Evidence from Chamorro*. Chicago and London: The University of Chicago Press.
- Chung, Sandra. 2003. The Syntax and Prosody of Weak Pronouns in Chamorro. *Linguistic Inquiry* 34: 547-599.
- Cohn, Abigail and John McCarthy. 1994. Alignment and Parallelism in Indonesian Phonology. *Working papers of the Cornell phonetics laboratory* 12: 53-137.
- Cohn, Abigail. 1992. The consequences of dissimilation in Sundanese. *Phonology* 9 :199-220.

- Cohn, Abigail. 1993. The initial dactyl effect in Indonesian and the status of degenerate feet. *Linguistic Inquiry* 24:372–381.
- Collins, Millard A., Virginia R. Collins and Sulfilix Hashim. 2001. *Mapun-English Dictionary*. Manila: Summer Institute of Linguistics.
- Condoravdi, Cleo and Paul Kiparsky. 2001. Clitics and clause structure. *Journal of Greek Linguistics* 2: 1-40.
- Condoravdi, Cleo and Paul Kiparsky. 2004. Clitics and clause structure: the Late Medieval Greek system. *Journal of Greek Linguistics* 5: 159–183.
- Cysouw, Michael. 2005. Morphology in the wrong place - a survey of preposed enclitics. ms. To appear in Dressler, Wolfgang U. (ed.) *Morphology and its Demarcations*. Amsterdam: Benjamins.
- DeGuzman, Videa P. 1988. Ergative analysis for Philippine languages: an analysis. In Richard McGinn (ed.), *Studies in Austronesian Linguistics. Southeast Asia series, no. 76*, 323-345. Athens, OH: Ohio University Center for International Studies.
- Demuth, K. 1995. Markedness and the development of prosodic structure. In Jill Beckman (ed.), *Proceedings of the North East Linguistic Society* 25: 13-25. Amherst, MA: GLSA, Univ. of Mass.
- Diesing, Molly, Dušica Filipović-Đurđević and Draga Zec. to appear. Clitic placement in Serbian: Corpus and experimental evidence. In Featherston, Sam and Susanne Winkler (eds.) *Proceedings from Linguistic Evidence 2008*. Mouton de Gruyter.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Dikken, Marcel den. 2006. *Relators and linkers: The syntax of predication, Predicate Inversion, and copulas*. Cambridge, MA: MIT Press.
- Dikken, Marcel. den. 2007. Phase Extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1-41.
- Dixon, R.M.W. and Alexandra Y. Aikhenvald. 2002. *Word: a cross-linguistic typology*. Cambridge: Cambridge University Press
- Dixon, Robert M. W. 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press
- Dost, Ascander and Vera Gribanova. 2006. Definiteness marking in the Bulgarian. In *Proceedings of WCCFL 25*, 132-140. Somerville: Cascadilla Press.

- Elfner, Emily. 2008. The interaction of linearization and prosody: Evidence from pronoun postposing in Irish and Scottish Gaelic. Ms., University of Massachusetts, Amherst.
- Embick, David and Rolf Noyer. 1999. Locality in post-syntactic operations. In V. Lin, C. Krause, B. Bruening, and K. Arregi (eds.), *Papers on morphology and syntax cycle two. MITWPL 34*.
- Embick, David, and Rolf Noyer. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32:555-595
- Emonds, Joseph. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.
- Engenhoven, Aone. 2004. *Leti: A Language of Southwest Maluku*. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 211. Leiden KITLV Press.
- Ernst, Thomas. 2001. *The Syntax of Adjuncts*. Cambridge Studies in Linguistics 96. Cambridge: Cambridge University Press.
- Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Ph.D. dissertation, MIT.
- Espinal, M. 1991. The representation of disjunct constituents. *Language* 67: 726-762.
- Esser, Samuel. 1934. *Handleiding voor de beoefening der Ledo-taal. Inleiding, teksten met vertaling en aantekeningen en woordenlijst*. Bandoeng: Nix.
- Everett, Daniel L. 1996. *Why there are no clitics: An alternative perspective on pronominal allomorphy*. SIL and the University of Texas at Arlington Publications in Linguistics, 123. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Fansler, Dean and Salvador Unson. 1916. Story of the Eventful Life of Princess Florentina of the Kingdom of Germany. *The Journal of American Folklore*, 29.112 :235-281.
- Fansler, Dean and Salvador Unson. 1916. Story of the Eventful Life of Princess Florentina in the Kingdom of Germany. *Journal of American Folklore* 29:235-281.
- Ferreirinho, Naomi. 1993. *Selected topics in the grammar of Limos Kalinga, the Philippines*. Pacific linguistics. Series B; no. 109.
- Fiengo, Robert & James Higginbotham. 1981. Opacity in NP. *Linguistic Analysis* 7, 395–421.

- Fikkert, Paula. 1994. *On the Acquisition of Prosodic Structure*. Ph.D. Dissertation , HIL dissertations 6, Leiden University.
- Fischer, Susann. 2002. *The Catalan Clitic System: A diachronic perspective on its syntax and phonology*. Berlin & New York: Mouton de Gruyter
- Fischer, Susann. 2003. Rethinking the Tobler-Mussafia Law: Data from Old Catalan Sources. *Diachronica* 20.2:259-288.
- Foley William A. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd LFG conference, Brisbane, Australia.
- Foley, William A. 2008. The place of Philippine languages in a typology Of voice systems. In Peter Austin and Simon Musgrave (eds.), *Voice and Grammatical Relations in Austronesian Languages*. Stanford: CSLI.
- Fontana, Josep. 1993. *Phrase Structure and the Syntax of Clitics in the History of Spanish*, Ph.D. dissertation, University of Pennsylvania.
- Fougeron, Cecile and Patricia A. Keating 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustic Society of America* 101, 3728-3740.
- Fox and Pesetsky, 2004. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31:1-46.
- Franks, Steven and Liljana Progovac. 1994. On the Placement of Serbo-Croatian Clitics. *Indiana Slavic Studies* 7, George Fowler, Henry R. Cooper, Jr., and Jonathan Z. Ludwig (eds.). Department of Slavic Languages and Literatures, Indiana University Bloomington. pp.1-10.
- Franks, Steven and Tracy-Holloway King. 2000. *A Handbook of Slavic clitics*. Oxford: Oxford University Press
- Franks, Steven. 1997. South Slavic Clitic Placement is *Still* Syntactic. *Penn Working Papers in Linguistics* 4.2:111-126.
- Franks, Steven. 1998. Clitics in Slavic. Position paper, Comparative Slavic Morphosyntax Workshop, Spencer, Indiana.
- Friberg, Barbara. 1988. Ergativity, focus and verb morphology in several South Sulawesi languages. In Roy Harlow & Ross Clark (eds.), *VICAL 2: Western Austronesian and contact languages: papers from the Fifth International Conference on Austronesian Linguistics*, 103-130. Auckland: Linguistic Society of New Zealand.

- Fukui, Naomi. 1986. *A theory of category projection and its applications*. Unpublished Ph.D. dissertation, MIT.
- Gelderen, Elly van. 2004. *Grammaticalization as Economy*. Amsterdam: John Benjamins.
- Gerds, Donna. 1988. Anti-passives and causatives in Ilokano: Evidence of an ergative analysis. In Richard McGinn (ed.), *Studies in Austronesian linguistics*, 295-321. Ohio University Monographs in International Studies, Southeast Asia Series No. 76. Athens, Ohio: Center for Southeast Asia Studies, Ohio University.
- Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's field trip, *CLS #7*, Chicago: University of Chicago, Chicago Linguistics Society.
- Grimshaw, Jane. 1991. *Extended Projections*. Ms. Brandeis University.
- Grimshaw, Jane. 1997. The Best Clitic: Constraint Conflict in Morphosyntax. In Lilianne Haegeman (ed.), *Elements of Grammar: Handbook of generative grammar*. Dordrecht: Kluwer.
- Grohmann, Kleanthes, J. Drury & J. C. Castillo. 2000. No more EPP. In R. Billerey and B.D. Lillehaugen (eds.), *WCCFL 19: Proceedings of the 19th West Coast Conference on Formal Linguistics*, 153-166. Somerville, Mass.: Cascadilla Press.
- Grosu, Alexander. 1981. *Approaches to Island Phenomena*. Amsterdam: North Holland.
- Gutiérrez Bravo, R. 2001. Phases, case and accessibility. In *Syntax and Semantics at Santa Cruz 3*. J. McCloskey (ed.). University of California, Santa Cruz.
- Hale, Ken. 1976. The adjoined relative clause in Australia. In R.M.W. Dixon (ed.) *Grammatical Categories in Australian Languages*, 78-105. Canberra: Australian Institute of Aboriginal Studies.
- Hale, Kenneth L. 1966. The Paman group of the Pama-Nyungan phyletic family. Appendix to *Languages of the World: Indo-Pacific Fascicle Six*, by G.N. O'Grady, C. F. & F.M. Voegelin. *Anthropological Linguistics* 8.2:162-197
- Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In K. Hale and J. Keyser (eds.), *The View from Building 20*. 111-176. MIT Press.
- Halpern, Aaron and Josep M. Fontana. 1994.  $X^0$  and  $X^{\max}$  clitics. In Eric Duncan, Donka Farkas, and Philip Spaelti (eds.) *Proceedings of WCCFL 12*. Mass.: Cascadilla Press.

- Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics*. Stanford: CSLI Publications.
- Hankamer, Jorge and Line Mikkelsen. 2005. When Movement Must be Blocked: A Reply to Embick and Noyer, *Linguistic Inquiry* 36:85-125.
- Harley, Heidi and Rolf Noyer. 1999. State-of-the-Article: Distributed Morphology. *Glott International* 4.4.
- Haverkort, Marco. 1994. Romance and Germanic clitics: A comparison of their syntactic behavior. In: *The Berkeley Conference on Dutch Linguistics*, T. Shannon and J. Snapper (eds.). Lanham MD: University Press of America.
- Hayes, Bruce. 1981. *A Metrical Theory of Stress Rules*, revised version of 1980 MIT doctoral dissertation, distributed by Indiana University Linguistics Club, Bloomington, Indiana.
- Hayes, Bruce. 1985. Iambic and trochaic rhythm in stress rules. In M. Niepokuj et al., (eds.). *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 429-446.
- Himmelman, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer Verlag.
- Himmelman, Nikolaus. 1999. The lack of zero anaphora and incipient person marking in Tagalog. *Oceanic linguistics* 38: 231-269.
- Himmelman, Nikolaus. 2008. Lexical categories and voice in Tagalog. In Peter K. Austin and Simon Musgrave (eds.), *Voice and Grammatical Functions in Austronesian Languages*. Stanford: CSLI.
- Hooper, Joan B. 1976. *Introduction to Natural Generative Phonology*. New York: Academic Press.
- Horn, G. 1974. *The Noun Phrase Constraint*, Ph.D. dissertation, University of Massachusetts, Amherst.
- Horst, J. van der. 1981. *Kleine Middelnederlandse syntaxis*. Amsterdam: Huis aan de drie grachten.
- Horwood, Graham. 2008. Implications of Katu nominalization for infixation typology. ms. Thammasat University. Available on the Rutgers Optimality Archive: <http://roa.rutgers.edu/files/969-0508/969-HORWOOD-0-0.PDF>

- Huck, Geoffrey J. and Younghee Na. 1990. Extraposition and Focus. *Language* 66.1:51-77.
- Hyman, Larry. 1977. On the nature of linguistic stress. In Larry Hyman (ed.), *Studies in Stress and accent*, (Southern California Occasional Papers in Linguistics 4), 37-82. Los Angeles: University of Southern California.
- Inkelas, Sharon and Draga Zec. 1988. Serbo-Croatian Pitch Accent: The Interaction of Tone, Stress and Intonation. *Language* 64:2.
- Inkelas, Sharon and Draga Zec. 1990. Prosodically Constrained Syntax. In Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*. Stanford: CSLI.
- Inkelas, Sharon, and Draga Zec. 1995. Syntax-phonology interface. In John Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, pp. 535-549.
- Inkelas, Sharon. 1990. *Prosodic Constituency in the Lexicon*. New York: Garland.
- Ito, Junko and Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*. 62-100. Cambridge, MA: Blackwell.
- Janse, Mark. 1998. Cappadocian Clitics and the Syntax-Morphology Interface. In Brian D. Joseph, Geoffrey C. Horrocks and Irene Philippaki-Warbuton (eds.), *Themes in Greek Linguistics II*, 257-281. Philadelphia: John Benjamins.
- Jelinek, Eloise and Richard Demers. 1994. Predicates and Pronominal Arguments in Straits Salish. *Language* 70:697-736.
- Jelinek, Eloise. 1996. Definiteness and Second Position Clitics in Straits Salish. In Aaron Halpern and Arnold Zwicky (eds.), *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI.
- John J. McCarthy. 2003. OT constraints are categorical. *Phonology* 20: 75-138
- Joseph, Brian. 1988. Pronominal Affixes in Modern Greek: The Case Against Clisis. *Proceedings of the Chicago Linguistic Society* 24, 203-215.
- Kahnemuyipour, Arsalan. 2003. Syntactic Categories and Persian Stress. *Natural Language and Linguistic Theory* 21: 333-379.
- Kaisse, Ellen. 1981. Luiseño particles and the universal behavior of clitics. *Linguistic Inquiry* 12:424-434.



- Kaisse, Ellen. 1982. Sentential clitics and Wackernagel's Law. *Proceedings of WCCFL 1*:1-14.
- Kaufman, Daniel. 2003. Paradigm effects and the shape/affix-generalization. In Gina Garding and Mimi Tsujimura (eds.), *Proceedings of WCCFL 22*,. Somerville, MA: Cascadilla Press.
- Kaufman, Daniel. 2005. Aspects of pragmatic focus in Tagalog. In I Wayan Arka and Malcolm Ross (eds.), *The Many Faces of Austronesian Voice Systems*. Canberra: ANU Press.
- Kaufman, Daniel. 2006. Rigidity versus relativity in adverbial syntax: evidence from Tagalog. In Hans-Martin Gärtner, Paul Law and Joachim Sabel (eds.), *Clause Structure and Adjuncts in Austronesian Languages*. Berlin: Mouton de Gruyter.
- Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: a Tagalog case study. *Theoretical Linguistics* 35.1:1-50.
- Kaufman, Daniel. to appear a. South Sulawesi pronominal clitics: form, function and position. *Studies in Philippine Languages and Cultures*.
- Kaufman, Daniel. to appear b. The morphosyntax of Maranao clitics. In Loren Billings & Nelleke Goudswaard (eds.), *Piakandatu ami Dr. Howard P. McKaughan..* Manila: SIL
- Kavitskaya, Darya. 2005. Prosodically conditioned fortition. Talk presented at the Manchester Phonology Meeting 13, May 26–28, 2005. Manchester, United Kingdom.
- Kayne, Richard S. 1969. *The Transformational Cycle in French Syntax*, Ph.D. dissertation, MIT.
- Kayne, Richard S. 1975. *French Syntax: The Transformational Cycle*, Cambridge, Ma: MIT Press.
- Kim, Yuni. 2008. *Topics in the Phonology and Morphology of San Francisco del Mar Huave*. Ph.D. dissertation, University of California, Berkeley.
- King, Tracy Holloway. 1996. Slavic clitics, long head movement, and prosodic inversion. *Journal of Slavic Linguistics* 4:274 – 311.
- Kiparksy, Paul. 1982. Lexical Phonology and Morphology. In In-Seok Yang (ed.), *Linguistics in the Morning Calm*. Seoul.

- Kiparsky, Paul. 2003. Finnish Noun Inflection. In Diane Nelson and Satu Manninen (eds.), *Generative Approaches to Finnic Linguistics*. Stanford: CSLI.
- Klavans, Judith L. 1994. *Clitics and Cliticization: The Interaction of Morphology, Phonology and Syntax*. Outstanding Dissertations in Linguistics Series. New York: Garland.
- Klavans, Judith. 1980. *Some Problems in a Theory of Clitics*. Ph.D. dissertation, University College London. [Distributed by the Indiana University Linguistics Club, Bloomington, 1982].
- Klavans, Judith. 1985. The Independence of Syntax and Phonology in Cliticization, *Language*, 61.1:95-120.
- Kramer, Ruth. to appear. The Amharic Definite Marker and the Syntax-Morphology Interface. *Syntax*.
- Kratzer, Angelika and Elisabeth Selkirk. 2007. Phase theory and prosodic spellout: the case of verbs. ms, to appear in *The Linguistic Review*.
- Kroch, Anthony. 1994. Morphosyntactic change. In *Proceedings of the 30<sup>th</sup> annual meeting of the Chicago Linguistics Society, Vol. 2*. Chicago Linguistics Society, Chicago, 180-201.
- Kroeger, Paul R. 1993. *Phrase structure and grammatical relations in Tagalog*. Dissertations in Linguistics. Stanford: CSLI .
- Kroeger, Paul R. 1998. Clitics and clause structure in Tagalog. In Maria Lourdes Bautista (ed.), *Pagtanaw*. Manila: DSLU Press.
- Kucerova, Ivana. 2005. The T-Extension Condition. In John Alderete et. al. (eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 227-235. Somerville, MA: Cascadilla Press.
- Laka, Itziar 1990. *Negation in syntax: On the nature of functional categories and projections*. Doctoral Dissertation, MIT, Cambridge, Massachusetts.
- Lamontagne, Greg. and Lisa Travis. 1986. The Case Filter and the ECP. *McGill Working Papers in Linguistics* 3: 51-75.
- Laughren, Mary. 2002. Syntactic constraints in a 'free word order language. In M. Amberber and P. Collins (eds.) *Language Universals and Variation*, 83-130. Westport Connecticut: Praeger Publishers.
- Lebeaux, David. 1988. *Language Acquisition and the Form of the Grammar*, Ph.D. dissertation, University of Massachusetts, Amherst.

- Lebeaux, David. 1990. Relative Clauses, Licensing, and the Nature of the Derivation. In *Proceedings of NELS 20*, GLSA, University of Massachusetts, Amherst, 318-332.
- Legate, Julie. 2002. *Warlpiri: Theoretical Implications*. Ph.D. dissertation, MIT.
- Legate, Julie. 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34.3.
- Legate, Julie. 2008. Warlpiri and the Theory of Second Position Clitics. *Natural Language and Linguistic Theory* 26.1.
- Legendre, Geraldine. 1998. Second position clitics in a V2 language: Conflict resolution in Macedonian. In: *Proceedings of the 1997 ESCOL Meeting* (Cornell University: CLC Publications), 139-149.
- Legendre, Geraldine. 2000. Morphological and prosodic alignment at work: The case of South-Slavic clitics. *Proceedings of the West Coast Conference on Formal Linguistics XVII*, 436-450.
- Legendre, Geraldine. 2001. Masked V2 effects and the linearization of functional features. In Geraldine Legendre, Sten Vikner, Jane Grimshaw (eds.), *OT Syntax*, 241-277. Cambridge: MIT Press.
- Liao, Hsiu-chuan. 2004. *Transitivity and Ergativity in Formosan and Philippine Languages*. Ph.D. dissertation, University of Hawai'i.
- Lieber, Rochelle. 1992. *Deconstructing morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.
- Lightfoot, David W. 1979. *Principles of diachronic syntax*. Cambridge Studies in Linguistics, 23. Cambridge: Cambridge University Press.
- Marantz, Alec. 1988. Clitics, Morphological Merger, and the Mapping to Phonological Structure. In Michael Hammond and Michael Noonan (eds.) *Theoretical Morphology*, 253-270. Ithaca: CLC Publications.
- Marantz, Alec. 1989. Clitics and Phrase Structure. In A. Kroch and M. Baltin (eds.), *Alternative Conceptions of Phrase Structure*, Chicago: University of Chicago Press, 99-116.
- Martinez, J. 1913. *Corrido at Buhay na Pinagdaanan ng Princesa Florentina sa Cahariang Alemaniam*. Manila: Imprenta de J. Martinez.

- Martins, Anna. 1995. A minimalist approach to clitic climbing. In *Papers from the 31st Regional Meeting of the Chicago Linguistic Society*, 215-239.
- McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation*, ed. Samuel Epstein and Daniel Seeley, 184–226. Oxford, Blackwell Publishers.
- Mercado, Raphael. 2002. Transparency is not less structure: A look at Tagalog Restructuring. Talk given at *CONSOLE 11*, Leiden.
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Mériz, Diane T. 1978. Observations on object collocation with finite verb parts in Medieval Occitan (to 1300). *Romania* 99 145-182.
- Miller, Philippi and Ivan Sag. 1997. French Clitics without Clitics or Movement. *Natural Language and Linguistic Theory* 15:573-639
- Miller, Phillip. 1992. *Clitics and constituents in Phrase Structure Grammar*. New York: Garland.
- Minde, Don van. 1997. *Malayu Ambong (Ambonese Malay): Phonology, Morphology, Syntax*. Leiden: Research School CNWS.
- Moreton, Elliott and Paul Smolensky. 2002. Typological consequences of local constraint conjunction. In Line Mikkelsen and Christopher Potts (eds.), *WCCFL 21 Proceedings*, Cambridge, MA: Cascadilla Press,
- Mussafia, Adolf. 1898. Enclisi o proclisi del pronome personale atono qualoggeto. *Romania* 27.105: 145–46.
- Naylor, Paz Buenaventura. 1980. Linking, relation marking and Tagalog syntax. In Paz Buenaventura Naylor (ed.), *Austronesian Studies*, Ann Arbor: University of Michigan. pp. 33-50.
- Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Nevis, Joel A. 1986. *Finnish particle clitics and general clitic theory*. Ohio State Univ. Working Papers in Linguistics 33.
- Noyer, Rolf. 1993. Mobile affixes in Huave: optimality and morphological well-formedness. In E. Duncan, D. Farkas and P. Spaelti (eds.), *Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, 67-82. Stanford: CSLI.

- Otani, Kazuyo. and John. Whitman 1991. V-raising and VP-ellipsis. *Linguistic Inquiry* 22:345-358.
- Pallesen, A. Kemp. 1985. *Culture contact and language convergence*. Linguistic Society of the Philippines, Special Monograph Issue, 24. Manila: Linguistic Society of the Philippines.
- Paster, Mary. 2006. *Phonological Conditions on Affixation*. Ph.D. dissertation, University of California, Berkeley.
- Pater, Joe. 1997. Metrical Parameter Missetting in Second Language Acquisition. In S.J. Hannahs and Martha Young-Scholten (eds.), *Focus on Phonological Acquisition*. Amsterdam: John Benjamins.
- Patricia A. Keating, Taehong Cho, Cecile Fougeron, and Chai shune Hsu. 2003. Domain-initial articulatory strengthening in four languages. *LabPhon*, 6:145-163.
- Penner, Zvi. 1991. Pronominal clitics in Bernese Swiss German and their structural position. Jakob Wackernagel and language acquisition. In Henk van Riemsdijk and Luigi Rizzi (eds.), *Clitics and their Hosts, ESF working papers*, 253-267.
- Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rhinehart & Winston.
- Petcova, Rossitza and Katya Alahverdzhieva. 2005. Transfer rules in Bulgarian-English Machine Translation: The POS (Noun) and the Noun Phrase. Presentation at the RoCoLi Summer School. Available on: [http://www.sfs.uni-tuebingen.de/rocoli/Ros\\_Kat\\_presentation.ppt](http://www.sfs.uni-tuebingen.de/rocoli/Ros_Kat_presentation.ppt)
- Peterson, David A. 2001. Ingush ?a: The elusive Type 5 clitic? *Language* 77: 144-55.
- Picallo, M. Carme. 1994. Catalan Possessive Pronouns: The Avoid Pronoun Principle Revisited. *Natural Language and Linguistic Theory* 12: 259-299.
- Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20: 365-424.
- Potts, Chris & Geoffrey K. Pullum. 2002. Model theory and the content of OT constraints. *Phonology* 19: 361-393.
- Predolac, Nikola. 2008. Against syntactic clitic placement in Serbian. ms. Cornell University.
- Prince, Alan. 1983. Relating to the Grid, *Linguistic Inquiry* 14:19-100.

- Progovac, Ljiljana. 1996. Clitics in Serbian/Croatian: comp as a the second position. In Aaron Halpern and Arnold Zwicky (eds.) *Approaching second*, 411-428. Stanford: CSLI.
- Progovac, Ljiljana. 2000. Where do clitics cluster? In Frits Beukema and Marcel den Dikken (eds.), *Clitic phenomena in European languages*. Amsterdam: John Benjamins.
- Quakenbush, J. Stephen and Edward Ruch. 2006. Pronoun ordering in Kalamianic. In *Proceedings of the Tenth International Conference on Austronesian Linguistics*. Manila: SIL. Available on: <http://www.sil.org/asia/philippines/ical/papers.html>
- Quakenbush, J. Stephen and Edward Ruch. 2006. Pronoun ordering in Kalamianic. *Tenth International Conference on Austronesian Linguistics*, 17-20 January 2006, Palawan, Philippines. Linguistic Society of the Philippines and SIL International. 19 p. <http://www.sil.org/asia/philippines/ical/papers.html>
- Quakenbush, J. Stephen. 2005. Some Agutaynen grammatical details: personal pronouns nominal markers, and tense/aspect/mode. In Danilo T. Dayag and J. Stephen Quakenbush (eds.) *Language and Education in the Philippines and beyond: A festschrift in honor of Maria Lourdes S. Bautista*, , 437-477. Manila: De La Salle University Press.
- Rackowski, Andrea and Norvin Richards. 2005. Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry* 36.4:565-599.
- Rackowski, Andrea. 2002. *The Structure of Tagalog: Specificity, Voice, and the Distribution of Arguments*. Ph.D. dissertation, MIT.
- Radanovic-Kocic, Vesna. 1996. The placement of SC clitics: A prosodic approach. In Aaron Halpern and Arnold Zwicky (eds.), *Approaching second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications.
- Ramos, Teresita V. 1971. *Tagalog structures*. Honolulu: Pacific and Asian Linguistics Institute. PALI language texts.
- Reid, Lawrence A. 1979. Towards a reconstruction of the pronominal systems of proto-Cordilleran, Phillipines. In Dang Nguyen Liem (ed.), *South-East Asian linguistic studies*, 259-275 (Pacific Linguistics, C 45). Canberra: Australian National University
- Reid, Lawrence A. 1981. Proto-Austronesian genitive determiners. In Andrew Gonzalez and David Thomas (eds.), *Linguistics across continents: Studies in honor of Richard S. Pittman*, 97-105. Linguistic Society of the Philippines Monograph, 2. Manila: SIL and Linguistic Society of the Philippines.

- Reid, Lawrence. 1982. The demise of Proto-Philippines. In Halim, Carrington, and Wurm (eds.), *Papers from the 3rd International Conference on Austronesian Linguistics, vol. 2: Tracking the travellers*. Pacific Linguistics, C-75.
- Richards, Norvin. 1999. Complementizer cliticization in Tagalog and English. *Toronto Working Papers in Linguistics 16.2: Proceedings of AFLA 6*, Carolyn Smallwood and Catherine Kitto (eds.). Toronto: Toronto Working Papers in Linguistics, 297-312.
- Richards, Norvin. 2003. Tagalog Ellipsis. *Cornell Working Papers in Linguistics 19: Proceedings of AFLA 9*, Anastasia Riehl and Thess Sevelia (eds.). Ithaca, NY: Cornell Working Papers in Linguistics, 228-251.
- Richards, Norvin. Against Bans on Lowering. *Linguistic Inquiry* 35.3.
- Richardson, Matthew. 1997. Czech clitics: 2P or not 2P, that is the question. Y.A.L.E. 1: Studies in the morphosyntax of clitics, L. Kaiser (ed.). pp. 131–150. Dept. of Linguistics, Yale University.
- Riemsdijk, H.C. van. 2005. Clitics: a state of the art report. In H.C. van Riemsdijk (ed.), *Clitics in the Languages of Europe. Language Typology, 8*, Berlin: Mouton de Gruyter.
- Rivero, María Luisa. 1986. Parameters in the typology of clitics in Romance and Old Spanish. *Language* 62: 774–807.
- Rivero, María Luisa. 1994. Clause Structure and V-Movement in the Languages of the Balkans. *Natural Language and Linguistic Theory* 12, 63-120.
- Rivero, Maria Luisa. 1997. On two locations for complement clitic pronouns: Serbo-Croatian, Bulgarian, and Old Spanish. In Ans van Kemenade and Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 170-206. Cambridge: Cambridge University Press.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.
- Rizzi, Luigi. 1986. On Chain Formation. In Hagit Borer, (ed.), *The Grammar of Pronominal Clitics*, Syntax and Semantics no. 19. Academic Press, New York.
- Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Lilianne Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer.
- Roberts, Ian. 1994. Two Types of Verb Movement in Romance. In Norbert Hornstein and David Lightfoot (eds.), *Verb Movement*. Cambridge: CUP.

- Rooth, Mats. 1985. *Association with focus*. Ph.D. dissertation. University of Massachusetts, Amherst.
- Ross, John R. 1967. *Constraints on variables in syntax*. Ph.D. dissertation, MIT.
- Rudin, C. and Kramer, C. and Billings, L. and Baerman. 1999. Macedonian and Bulgarian li questions: Beyond syntax, *Natural Language and Linguistic Theory* 17.3:541-586.
- Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6:445-501.
- Sadock, J.M. 1991. *Autolexical syntax. A theory of parallel grammatical representations*. [Studies in Contemporary Linguistics] Chicago: The University of Chicago Press.
- Sadock, Jerrold. 1995. A multihierarchy view of clitics. *Proceedings of the Chicago Linguistic Society* 31, part 2. 258-79.
- Sadock, Jerrold. 2003. *A grammar of Kalaallisut (West Greenlandic Inuktitut)*. Munchen: Lincom Europa
- Sag, Ivan, Thomas Wasow and Emily Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI.
- Schachter, Paul and Fe T. Otones. 1972. *Tagalog Reference Grammar*. Berkeley: California University Press.
- Schachter, Paul. 1973. Constraints on clitic order in Tagalog. In: *Parangal kay Cecilio Lopez, A. B. Gonzalez* (ed.). pp. 214–231. Manila: Linguistic Society of the Philippines.
- Schachter, Paul. 1976. The subject in Philippine languages: topic, actor, actor-topic, or none of the above. In *Subject and Topic*, C. Li, ed. New York: Academic Press.
- Schachter, Paul. 1996. *The subject in Tagalog: still none of the above*. UCLA Occasional Papers in Linguistics 15. Los Angeles: UCLA Department of Linguistics.
- Schütze, Carson. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. In Andrew Carnie, Heidi Harley and Tony Bures (eds.), *Papers on phonology and morphology*. MIT Working Papers in Linguistics 21, 373-473.



- Selkirk, Elisabeth. 1980. The role of prosodic categories in English word stress. *Linguistic Inquiry* 11.563-606
- Selkirk, Elisabeth. 1995. The prosodic structure of function words. In Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, 439-70. Amherst, MA: GLSA Publications.
- Selkirk, Elisabeth. 1996. Sentence prosody: Intonation, stress, and phrasing. In John Goldsmith (ed.), *The Handbook of Phonological Theory*, chap. 16, Blackwell, London, pp. 550–569.
- Selkirk, Elizabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press
- Shaer, Benjamin. 2003. An "Orphan" Analysis of Long and Short Adjunct Movement in English. In Gina Garding and Mimura Tsujimura (eds.), *Proceedings of WCCFL XXII*, 450-463. Cascadilla Press.
- Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
- Simpson, Jane Helen. 1991. *Warlpiri morpho-syntax: a Lexicalist Approach*. Dordrecht: Kluwer Academic.
- Sityar, Emily M. 1989. *Pronominal Clitics in Tagalog*. M.A. thesis, University of California, Los Angeles
- Smith, Ian, and Steve Johnson. 2000. Kugu Nganhcara. In R. M. W. Dixon and B. Blake (eds.), *The Handbook of Australian Languages, Volume 5*, 355–489. Melbourne: Oxford University Press Australia.
- Smith, Jennifer. 2002. *Phonological Augmentation in Prominent Positions*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Paper presented at Rutgers Optimality Workshop 1.
- Sneddon, James Neil. 1975. *Tondano Phonology and Grammar*. Pacific Linguistics; B 38. Canberra: Australian National University.
- Steedman, Mark. 2000. *The Syntactic Process*. Cambridge: MIT Press.
- Steven, Franks. 2000. Clitics at the Interface. In Fritz Beukema and Marcel den Dikken (eds.), *Clitic Systems in European Languages*, 1-46.

- Suñer, Margarita. 1975. Spanish adverbs: Support for the phonological cycle? *Linguistic Inquiry* 6.4:602–605.
- Svenonius, Peter. 2004. On the Edge. In David Adger, Cécile de Cat, and George Tsoulas (eds.), *Peripheries: Syntactic Edges and their Effects*, pp. 261-287. Dordrecht: Kluwer
- Tawasil, Sapii. 1978. In kahalan sin hi kuting iban hi ambaw. *Studies in Philippine Linguistics* 2.2:192-201.
- Taylor, Ann. 1990. *Clitics and Configurationality in Ancient Greek*. Ph.D. dissertation, University of Pennsylvania, Philadelphia.
- Taylor, Ann. 1996. A prosodic analysis of clitic position in Ancient Greek. In Aaron Halpern and Arnold Zwicky (eds.), *Second Position Clitics and Related Phenomena*. Stanford: CSLI.
- Terzi, Arhonto. 1999. Clitic combinations, their hosts and their ordering. *Natural Language and Linguistic Theory* 17:85-121.
- Tobler, Adolf. 1889. Pronominales Objekt zu Infinitiv oder Partizipium. *Zeitschrift für romanische Philologie* 13:186-191. Reprinted in: *Vermischte Beiträge zur Französischen Grammatik* 2. 1902 and 1971 [Amsterdam: Rodopi]
- Tomic, Olga Mišeška. 1996a. The Balkan Slavic Clausal Clitics. *Natural language and Linguistic Theory* 14:811–872.
- Tomić, Olga Mišeška. 1996b. The Balkan Slavic Nominal Clitics. In Aaron L. Halpern and Arnold M. Zwicky (eds.), *Approaching Second: Second position clitics and Related Phenomena*, 511-536. Stanford: CSLI Publications.
- Travis, Lisa and Andrea Rackowski. 2000. V-initial languages: X or XP movement and adverbial placement. In A. Carnie and E. Guilfoyle (eds.), *The Syntax of Verb-initial languages*. Oxford: Oxford University Press.
- Truckenbrodt, Hubert. 1995. *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. Ph.D. dissertation, MIT.
- Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30: 219-255.
- Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26: 79-123.

- van Riemsdijk, Henk C. 1999. Clitics: A state of the art report. In Henk van Riemsdijk (ed.), *Clitics in the Languages of Europe*. Berlin: Mouton de Gruyter.
- van Riemsdijk, Henk. 1990. Functional prepositions. In H. Pinkster and I. Gene (eds.), *Unity in Diversity*, 229–241. Dordrecht: Foris.
- Venneman, Theo. 1988. *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin: Mouton de Gruyter.
- Wackernagel, Jacob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1.333-436.
- Wagner, Michael. 2005. *Prosody and Recursion*. Ph.D. dissertation. MIT.
- Wanner, Dieter. 1987. *The Development of Romance Clitic Pronouns from Latin to Old Romance*. (Empirical Approaches to Language Typology, 3), Berlin: Mouton-de Gruyter.
- Wanner, Dieter. 1996. Second position clitics in Medieval Romance. In Aaron L. Halpern and Arnold M. Zwicky (eds.), *Approaching Second: Second position clitics and related phenomena*, 537-579. Stanford: CSLI.
- Wegmüller, Ursula. 1993. *Klitika im Vicentino*. Universität Bern, Arbeitspapier 31.
- Werle, Adam. 2004. Enclisis and proclisis in Serbian/Croatian. Talk given at FASL 13, University of South Carolina.
- Werle, Adam. 2009. *Word, phrase, and clitic prosody in Bosnian, Serbian, and Croatian*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Whitman, John. 1997. Kakarimusubi from a Comparative Perspective. In J. Haig and H. Sohn (eds.), *Japanese Korean Linguistics* 6, 161-178. Stanford: CSLI.
- Wilder, Chris and Damir Čavar. 1994. Word Order Variation, Verb Movement and Economy Principles. *Studia Linguistica* 48.1:46-86.
- Wolff, John, Maria Theresa C. Centeno & Der-Hwa V. Rau. 1991. *Pilipino Through Self-Instruction*. Four volumes. SEAP, Cornell University.
- Wolff, John. 1967. History of the Dialect of the Camotes Islands, Philippines, and the Spread of Cebuano Bisayan. *Oceanic Linguistics* 6:63-79.
- Wurmbrand, Susanne. 2003. *Infinitives: Restructuring and Clause Structure*. Berlin: Mouton de Gruyter.

- Yu, Alan C. L. 2007. *A Natural History of Infixation* (Oxford Studies in Theoretical Linguistics 15). Oxford: Oxford University Press.
- Zagona, Karen. 1990. Mente adverbs, compound interpretation and the projection principle. *Probus* 2:1–30.
- Zec, Draga & Sharon Inkelas. 1990. Prosodically constrained syntax. In Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 365-378. Chicago and Stanford: University of Chicago Press and CSLI Publications.
- Zec, Draga. 2005. Prosodic Differences among Function Words. *Phonology* 22:77-112.
- Zorc, David. 1975/1977. The Bisayan dialects of the Philippines: Subgrouping and reconstruction. (Pacific Linguistics C-44) Canberra. [Ph.D. dissertation, Cornell University.]
- Zorc, R. David. 1986. The genetic relationships of Philippine languages. In Paul Geraghty et al. (eds.) *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics*, Pacific Linguistics C-94. Canberra. 147-173.
- Zribi-Hertz & Mbolatianavolana .1999. Towards a modular theory of linguistic deficiency: Evidence from Malagasy personal pronouns. *Natural Language and Linguistic Theory*. 17.1:161-218.
- Zubizarreta, Maria Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge, MA: MIT Press.
- Zwicky & Pullum 1989 Misconceiving morphology: A discussion of Bromberger & Halle's "Conceptual issues in Morphology". Ms, Standord University. Presented at the CSLI Conference on the Foundations of Phonology.
- Zwicky, Arnold and Geoffrey Pullum. 1983. Cliticization versus inflection: English n't. *Language* 59, 502-513.
- Zwicky, Arnold And Geoffrey Pullum. 1986. Two spurious counterexamples to the principle of phonology-free syntax. *Ohio State University Working Papers in Linguistics* 32: 92-99
- Zwicky, Arnold. 1977. *On Clitics*. Indiana University Linguistics Club.
- Zwicky, Arnold. 1985. Clitics and particles. *Language* 61:283-305.
- Zwicky, Arnold. 1994. Clitics. In R. A. Asher and J. M. Y Simpson (eds.), *The Encyclopedia of Language and Linguistics*, 571–576. Oxford: Pergamon Press.