

Proceedings of the Twenty-sixth
**Western Conference
On Linguistics**

Volume Nine

WECOL 96

Held at
The University of California, Santa Cruz
October 96

Edited by
Vida Samiian

Department of Linguistics
California State University, Fresno
Fresno, California, 93740-0092

- Copyright © 1998 by the
Department of Linguistics
California State University, Fresno
- "Asymmetric Head Movement and
Parsing" Copyright © 1998 by
Peter Ackema and Ad Neeleman
- "Mostly Predictable: Cyclicity and the
Distribution of Schwa in Itelmen"
Copyright © 1998 by Jonathan
David Bobaljik
- "Relations Between Core and Periphery
in American Sign Language
Phonology" Copyright © 1998 by
Diane Brentari
- "The Role of the Determiner in the
Semantic Interpretation of NP-type
Nouns" Copyright © 1998 by
Tracy C. Burns and Nancy N. Soja
- "Optimizing Structure in Context:
The Case of German Scrambling"
Copyright © 1998 by Hye-Won Choi
- "Explaining Coronal Underspecification"
Copyright © 1998 by Michael
Dobrovolsky
- "Exceptionality in Finnish"
Copyright © 1998 by Andrew Dolbey
and Ronald Sprouse
- "Restricting Conjunction to Constraint
Families" Copyright © 1998 by
Haruka Fukazawa and Viola Miglio
- "Phonological Underspecification and
the Subset Principle" Copyright ©
1998 by Mark Hale and Charles Reiss
- "The Role of Comprehension,
Reinterpretation and the Uniformity
Condition in Historical Change: The
Case of the Development of Cl Clusters
from Latin to Hispanico-Romance"
Copyright © 1998 by D. Eric Holt
- "Linearization of Verb Clusters in West
Germanic" Copyright © 1998 by
Andreas Kathol
- "Emergence of the Markedness
Constraint No-Lar and its Split
Behavior in the Reduplicative
Domain" Copyright © 1998 by
Soohee June Kim
- "Two Types of Bare Measure Phrases"
Copyright © 1998 by Murat Kural
- "Focus in the Future and the
Thetic/Categorical Distinction"
Copyright © 1998 by Felicia A. Lee
- "Adverbial Quantifiers, Maximal
Situations, and "Weak" E-type
Pronouns" Copyright © 1998 by
Toshiyuki Ogihara
- "When the Best Isn't Good Enough:
Phonologically-conditioned
Ungrammaticality in Optimality
Theory" Copyright © 1998 by
Cemil Orhan Orgun and Ronald
Sprouse
- "Durational Contrasts and the
Iambic/Trochaic Law" Copyright ©
1998 Anthi Revithiadou and Ruben
van de Vijver
- "Subjacency Forever" Copyright ©
1998 by Norvin Richards
- "Evidence for the Optional Tense
Hypothesis: Tense in Subordinate
Clauses in the Acquisition of
English" Copyright © 1998 by
Taylor Roberts
- "On 'Fewest Steps'" Copyright © 1998
by E.G. Ruys
- "English Nominals without Overt Noun
Heads: A Lexical Sharing
Analysis" Copyright © 1998 by
Michael T. Westcoat

ISBN 1-879890-05-4

Contents

Asymmetric Head Movement and Parsing	1
<i>Peter Ackema and Ad Neeleman</i>	
Mostly Predictable: Cyclicity and the Distribution of Schwa in Itelmen	14
<i>Jonathan David Bobaljik</i>	
Relations Between Core and Periphery in American Sign Language Phonology	29
<i>Diane Brentari</i>	
The Role of the Determiner in the Semantic Interpretation of NP-type Nouns	43
<i>Tracy C. Burns and Nancy N. Soja</i>	
Optimizing Structure in Context: The Case of German Scrambling	56
<i>Hye-Won Choi</i>	
Explaining Coronal Underspecification	72
<i>Michael Dobrovolsky</i>	
Exceptionality in Finnish	85
<i>Andrew Dolbey and Ronald Sprouse</i>	
Restricting Conjunction to Constraint Families	102
<i>Haruka Fukazawa and Viola Miglio</i>	
Phonological Underspecification and the Subset Principle	118
<i>Mark Hale and Charles Reiss</i>	

The Role of Comprehension, Reinterpretation and the Uniformity Condition in Historical Change: The Case of the Development of Cl Clusters from Latin to Hispano-Romance	133
<i>D. Eric Holt</i>	
Linearization of Verb Clusters in West Germanic	149
<i>Andreas Kathol</i>	
Emergence of the Markedness Constraint No-Lar and its Split Behavior in the Reduplicative Domain	162
<i>Soohee June Kim</i>	
Two Types of Bare Measure Phrases	177
<i>Murat Kural</i>	
Focus in the Future and the Thetic/Categorical Distinction	188
<i>Felicia A. Lee</i>	
Adverbial Quantifiers, Maximal Situations, and "Weak" E-type Pronouns	200
<i>Toshiyuki Ogihara</i>	
When the Best Isn't Good Enough: Phonologically-conditioned Ungrammaticality in Optimality Theory	214
<i>Cemil Orhan Orgun and Ronald Sprouse</i>	
Durational Contrasts and the Iambic/Trochaic Law	229
<i>Anthi Revithiadou and Ruben van de Vijver</i>	
Subjacency Forever	243
<i>Norvin Richards</i>	

Evidence for the Optional Tense Hypothesis: Tense in Subordinate Clauses in the Acquisition of English	256
<i>Taylor Roberts</i>	
On "Fewest Steps"	269
<i>E.G. Ruys</i>	
English Nominals without Overt Noun Heads: A Lexical Sharing Analysis	284
<i>Michael T. Wescoat</i>	

ASYMMETRIC HEAD MOVEMENT AND PARSING*

Peter Ackema & Ad Neeleman

Utrecht University

1. INTRODUCTION

Kayne (1994) proposes a theory in which linear order is determined by c-command relations in the syntactic tree. His theory is built on the axiom that if a node A c-commands a node B, then the terminals dominated by A must precede the terminals dominated by B. Mutual c-command thus leads to contradictory demands on the order of terminals and is consequently excluded.

This theory has a number of interesting empirical consequences. From the principles proposed by Kayne it follows, for example, that rightward movement cannot exist. Movement must be to a c-commanding position in order for the trace to be licensed, and, as explained above, c-commanding positions will be found only to the left of the origin of movement. The exclusion of rightward movement appears to be an attractive outcome, since several well-motivated movements are indeed invariably leftward.

Another important consequence of Kayne's theory is that syntactic structures must be right-branching: left-branching trees go 'upward' from left to right, instead of 'downward'. One result that Kayne obtains from this is a uniform underlying order, in which specifiers precede their heads while complements follow it. Variation in surface order must then be the consequence of movement. This particular consequence of Kayne's proposals faces some empirical problems. Both within and across languages, the order of elements preceding a head is often the reverse of the order that can be observed when the same elements follow the head. This necessitates a different view of base-generated structures: X-bar theory is symmetric in that it allows a constituent to be attached both to the right and to the left of the head.

This leaves unexplained the asymmetries between leftward and rightward movement that motivated Kayne's theory. We propose that, as a result of some independently motivated principles of the human parser, movement to the right is severely restricted (but crucially not ruled out across the board). We will restrict ourselves to head movement in this paper.

2. SYMMETRIC SYNTAX

Traditional X-bar theory allows complements, specifiers and adjuncts to be attached both to the left and to the right of the head in principle. Given that constituents are attached at a specific X-bar level, such a theory predicts mirror image effects: the order in which two base-generated elements appear to the left of the head is the reverse of the order of those elements if they are generated to the right of the head:

- (1) [XP₁ [XP₂ [Y] XP₂] XP₁]

Due to the condition that DP arguments must be assigned case in a specific direction and sometimes also under adjacency, arguments in many languages cannot be generated on both sides of the head.¹ For elements within VP that *can* be generated on both sides of the head, however, mirror image effects can indeed be observed. Their distribution confirms the view that base generation is symmetric.

A well-known example is the mirror image effect that occurs with PPs in modern Dutch, see (2) (see Koster 1974).

- (2) a. dat Jan [[tijdens de pauze]₁ [[aan zijn vader]₂ dacht]]
that John during the break of his father thought
 a'. ??dat Jan [[aan zijn vader]₂ [[tijdens de pauze]₁ dacht]]
 b. dat Jan [[dacht [aan zijn vader]₂] [tijdens de pauze]₁]
 b'. ??dat Jan [[dacht [tijdens de pauze]₁] [aan zijn vader]₂]]

A further example of mirror image effects can be found in Middle Dutch where NP arguments and resultatives could be generated on both sides of the verb, as in (3) (see Neeleman & Weerman 1992). Irrespective of their position with respect to the verb, the resultative appears closer to V than the object.

- (3) a. So suldy [[den coeke]₁ [[in stucken]₂ wryven]]
thus should-you the cake to pieces rub
 b. Si hebben [[gevaerwet [root]₂] [die straten]₁]
they have painted red the streets

Mirror image effects can not only be observed within one language, but also cross-linguistically. The order of preverbal adverbials in Dutch, for instance, turns out to be the mirror image of the order of postverbal adverbials in English, see (4).

- (4) a. dat Jan [[gisteren]₁ [[vurig]₂ [een meisje kuste]]]
that Jan yesterday passionately a girl kissed
 a'. ?*dat Jan [[vurig]₂ [[gisteren]₁ [een meisje kuste]]]
 b. John [[[kissed a girl] [passionately]₂ [yesterday]₁]
 b'. ?*John [[[kissed a girl] [yesterday]₁] [passionately]₂]

Mirror image effects do not only occur in the verbal, but also in the nominal domain. A language-internal example of this phenomenon can be found in Tagalog (Norvin Richards, p.c.). Tagalog adjectives can appear on either side of the noun. In those cases where two nonconjoined adjectives are generated on the same side of the noun, their order shows a mirror-image effect:

- (5) a. pinakamalapit na pulang bahay
nearest LINK red-LINK house
 a'. *pulang pinakamalapit na bahay
 b. bahay na pulang pinakamalapit
 b'. *bahay na pinakamalapit na pula

In fact, the argument holds more generally. Greenberg (1966) observes that in the majority of cases the order of determiners, numerals and adjectives in languages in which these elements follow the noun is the mirror image of the order found in languages in which they precede the noun.

- (6) a. determiner - numeral - adjective - noun
 b. noun - adjective - numeral - determiner

In conclusion, a symmetric X-bar theory predicts that elements generated to the right of the head appear in the reverse order of elements generated to its left, if no movement occurs. The data discussed in this section appear to confirm this prediction.

3. ASYMMETRIC PARSING

In addition to the mirror image effects mentioned in section 2, some specific anti-mirror image effects are attested. In a number of noun-initial languages the dependent elements show up in the same order as in noun-final languages, see (7a). This can be explained by head movement of the noun to the left, starting with the basic structure in (6a). An asymmetry can now be observed: rightward head movement of the noun is

impossible. Alongside (6b), no languages exist in which the order in (7b) is found (Greenberg 1966).

- (7) a. noun - determiner - numeral - adjective
 b. *adjective - numeral - determiner - noun

The asymmetry of head movement is also attested in the verbal domain. The assumption that subjects can be generated on either side of V' and objects on either side of V gives rise to an acceptable language typology. In addition to this, some languages display head movement of V to the left. Examples are VSO languages and languages with Verb Second. There are no languages, however, in which there is straightforward distributional evidence for rightward verb movement.

If syntax is not fundamentally asymmetric, we must either assume a principle of asymmetry specific to movement, or look for the explanation of the observed asymmetries elsewhere. We pursue the latter option here. In particular, we suggest an explanation in terms of universal parsing strategies.

From the perspective of the parser, leftward movement and rightward movement differ in that rightward movement requires the introduction of a trace in a (partly) analyzed string, whereas leftward movement allows the trace to be introduced at the same time that the string is analyzed. Introduction of a trace in a (partly) analyzed string sometimes necessitates destruction of already established information, which is impossible. Let us now consider the properties of the parser in more detail.

Due to the temporal order in which the input string is received, the parser scans this string from left to right, building up a representation of the sentence as it goes along. The representation that forms the output of the parser consists of a set of assertions about a syntactic tree (cf. Marcus et al. 1983). At first, the information that can be established is rather crude. Only a rough outline of the tree can be given, in which precedence and dominance (though not immediate dominance) relations are noted. As the parser proceeds, the set of assertions becomes more precise, so that toward the end of the parsing process an accurate description of the structure results. This refinement of the set of assertions is subject to a condition of informational monotonicity, as Berwick & Weinberg (1985) call it. Information can be added to the assertion set, but no already established information can be altered.

Parsing seems to be incremental. All decisions that the parser can make on the basis of the current input symbol and the left context are made immediately. A consequence of this is early attachment. The parser tries to attach analyzed constituents to a projection line as soon as possible, even in case it has not encountered the head of this projection line yet (Frazier 1987). What we assume, therefore, is that, in addition to the information mentioned above, the parser immediately notes whether a constituent forms a right branch or left branch of a projection.

Incremental parsing further implies that the parser has no lookahead (cf. Frazier & Rayner 1982, Gorrell 1995). This means that at any stage, the parser can only consider the current input symbol and (a limited amount of) the already parsed material when it is deciding what to do next. Crucially, information from its right context is not available.

This leaves open the problem of what the parser does when it encounters input that is locally ambiguous. Our analysis of how rightward movement can be parsed does not hinge much on the solution adopted for this problem. What we assume here is that if there are two analyses that comply with the pertinent grammar, the parser notes both, and pursues the one that is favored by some strategies defining the preferred parse. If this parse turns out to be incompatible with newly received material, it is aborted and the next best option will be pursued.

However, as a consequence of informational monotonicity such backtracking is only possible if the parser can detect the local ambiguity immediately on the basis of the current input symbol and the left context. If at some point P the parser cannot decide between analyses, it may first try one of the analyses and in case of failure try the other. But if at P the input string allows only one analysis, each continuation of the parse after P must comply with this analysis. If material after point P is incompatible with it, the sentence is not parsable.

Finally, we assume that the parser handles antecedent-gap relations by applying a filler-driven strategy (cf. Frazier 1987, 1993): the postulation of a gap depends on the presence of an antecedent. Once an antecedent is identified as such, a position to insert a gap in is looked for.

Summarizing, the following parsing instructions hold:

- (8)
 - a. Construct an assertion set which describes a tree, scanning the input string from left to right and using no lookahead.
 - b. Do not postulate a trace without having encountered an antecedent.
 - c. Do not destroy already established information.

- d. If the input string allows two or more grammatical analyses, pursue the one that adheres best to the preference strategies.
- e. If the present analysis is incompatible with grammatical principles, abort it and proceed to the one that adheres next best to the preference strategies.

We will argue that the problem with most instances of head movement to the right concerns (8c), because such movement can have the effect that information must be added to the parse that conflicts with already established information. Moreover, we will show that exactly when this is not the case, head movement to the right is possible.

4. INFELICITOUS RIGHTWARD HEAD MOVEMENT

Consider the effects of the principles in (8). Suppose that the first part of an input string has been analyzed as an XP.² The parser will mark this XP as a left branch (LB). It could only be marked as a right branch (RB) if a trace is introduced by the parser first, but because of (8b) this trace cannot be motivated then: no antecedent has been identified yet.

(9) XP_{LB}

The parser will continue to describe a right-branching tree (that is, it marks every XP as occupying a left branch) until it encounters the head of the projection hosting the already analyzed XPs:

(10) $XP_{LB} XP_{LB} XP_{LB} Y$

At this point there are two options. In the first option, the head has not been moved. Consequently, the tree becomes left-branching after the head, given the symmetry of basic structures (see (1)):

(11) $XP_{LB} XP_{LB} XP_{LB} Y XP_{RB} XP_{RB} XP_{RB}$

The second option after the parser has encountered the head is to assume that it has been moved. The branching direction must then remain constant, since the moved head would not c-command its trace otherwise. The consequence is that the XPs following the moved head will be put on a left branch until the trace of the head is inserted. This trace indicates the base position of the head, and therefore the branching direction must

change after it. One possible result is given in (12) (with the trace in penultimate position).

(12) $XP_{LB} XP_{LB} XP_{LB} Y_i XP_{LB} XP_{LB} t_i XP_{RB}$

Of the options illustrated in (11) and (12), the parser will pursue the one that complies with the principles of the pertinent grammar. If both options comply with that grammar, it will pursue the preferred one first.

So, the parser can analyze the string under discussion either as a basic structure or as a structure involving head movement to the left. However, an analysis in which the head has been moved to the right over one or more of the XPs will never be postulated. Due to (8b), a trace cannot be inserted without proper motivation, that is, a trace can only be inserted after the parser has encountered the antecedent. In the case of head movement to the right, this has the effect that the trace of Y must be inserted in a position preceding one or more XPs already put on a left branch. However, this implies that these XPs should in fact be placed on a right branch. In other words, information already contained in the parse must be altered, which is impossible according to (8c). So, (13b), for example, is not a possible continuation of the parse in (13a).

(13) a. $XP_{LB} XP_{LB} XP_{LB}$
 b. $XP_{LB} XP_{LB} t_i XP_{RB} Y_i$

In the case of head movement to the left this problem does not arise. Since the antecedent precedes its trace, this trace can be introduced at the moment that the branching direction is determined.

This explanation is not undermined by the property of the parser that it may postulate two alternatives at a locally ambiguous point, and can backtrack if the one pursued first fails. Because of the filler-driven strategy to gap postulation, the parser never faces a local ambiguity at the point where the trace of a rightward-moved head should be inserted. When encountering XPs that precede the head it simply must put them on a left branch.

5. FELICITOUS RIGHTWARD HEAD MOVEMENT

The argumentation in section 4 implies that movement to the right is possible if the LB/RB specifications do not have to be altered after introduction of the trace. This means that for the string discussed above there are in fact two additional possibilities: a parse with XP-movement

to the right, for instance (14a), and a parse with string-vacuous head movement to the right, as in (14b).³

- (14) a. $XP_{LB} XP_{LB} XP_{LB} Y XP_{RB} t_i XP_{RB} XP_{i,RB}$
 b. $XP_{LB} XP_{LB} XP_{LB} t_i Y_i XP_{RB} XP_{RB} XP_{RB}$

In this section we will show that if no LB/RB specifications need to be changed, rightward head movement is indeed possible (for rightward XP-movement, see Rochemont 1990).

The first type of rightward head movement we will discuss concerns stranded prepositions in Dutch. It can be argued that these incorporate into the verb (cf. Sturm & Kerstens 1978 and Hoeksema 1991). The observation on which this claim is based is that stranded prepositions in Dutch must be adjacent to the verb:

- (15) a. Daar_i wil ik tijdens de lunch [t_i t_j] [over_j praten]
that want I during the lunch about speak
 b. *Daar_i wil ik [t_i over] tijdens de lunch praten
that want I about during the lunch speak

This restriction cannot be reduced to the distribution of PPs in general, since PPs do occur in positions preceding adverbials, see (16a). It can also not be reduced to a restriction on subextraction from constituents that are not adjacent to the verb, since this is allowed in constructions like (16b).

- (16) a. Ik wil daarover tijdens de lunch praten
I want that-about during the lunch speak
 b. Wat_i heb je [t_i voor mensen] tijdens de lunch gezien?
what have you for people during the lunch seen

The example of P-incorporation in (15a) is a case of string-vacuous rightward head movement, parsable on a par with (14b). It is remarkable, however, that in other cases the movement may be nonstring-vacuous: the preposition may cross a dependent of the verb. The relevant examples are given in (17).

- (17) a. dat ik de deur daarmee groen verf
that I the door that-with green paint

- b. *dat ik de deur groen daarmee verf
that I the door green that-with paint
- c. de verf waar_i ik de deur [t_i t_j] [_V mee_j [_V groen verf]]
the paint that I the door with green paint
- d. de verf waar_i ik de deur [t_i t_j] [_V groen [_V mee_j verf]]
the paint that I the door green with paint

There is reason to believe that *groen verven* 'green paint' is a complex predicate, that is, a complex V^O category, generated by adjunction of the resultative to the verb (see Neeleman 1994). This immediately accounts for (17c): although the stranded preposition is not adjacent to the verb, it is adjacent to the verbal complex. It has incorporated into the higher V^O node of the complex predicate. Crucially, the lower V^O segment is also a target for P-incorporation, which accounts for (17d).

What (17d) shows is that rightward head movement may cross a dependent of the head into which incorporation takes place. That this is possible follows from the parsing strategies outlined above. Consider how the parser analyzes a string corresponding to the construction under discussion. As before, XPs preceding the lexical head are put on a left branch (see 18a). When the parser encounters the preposition and hypothesizes that it is an antecedent (and given the grammar of Dutch it must), it may introduce a trace in a position before the resultative, since this does not mean that established assertions about the structure must be altered. The resultative can remain on a left branch because it is not in the projection headed by the trace, but in the one headed by the verb. So, (18b) is a possible continuation of (18a).

- (18) a. $XP_{LB} XP_{LB} XP_{LB}$
 b. $XP_{LB} XP_{LB} t_i XP_{LB} P_i V$

There is, in other words, no absolute adjacency condition on rightward incorporation. Given this, it is surprising that rightward incorporation may not cross dependents of the incorporating head itself. This can be shown with a construction that involves nonstring-vacuous V-movement to the right, namely Dutch V-to-V raising, as in (19) (cf. Evers 1975).

- (19) dat Jan [Marie de samba t_i] [zag dansen_i]
that Jan Marie the samba saw dance

From the perspective of the parser, (19) involves a number of substrings to be analyzed as XPs, followed by two verbs. This string can be analyzed by the parser as in (20). The trace of the final verb is inserted after the parser has encountered this verb. Since the trace is inserted in a position following all XPs no changing of LB/RB specification is required. The parse in (20) agrees with Evers' analysis of V-to-V raising (although further analysis is of course required to arrive at (19)).

(20) $XP_{LB} XP_{LB} XP_{LB} t_i V V_i$

Although we have shown that there is no adjacency condition on rightward head movement, the embedded verb cannot be raised across one of its own dependents.⁴ This can be argued on the basis of an observation by Reuland (1990). In Dutch, a postverbal PP can take scope over a preverbal adverbial (as expected given the symmetry of basic structures). The example in (21) can mean both that it was the case for some time that Jan frequently hampered the project or that it was frequently the case that Jan hampered the project for some time.

(21) *dat Jan het project regelmatig hinderde gedurende een tijdje*
that Jan the project frequently hampered for some time

Suppose the verb in (21) adjoins to a higher verb, across the postverbal PP. We would then still expect the same ambiguity. However, in a V-to-V raising construction like (22) only one reading is available, namely the reading in which the adverbial takes scope over the PP.

(22) *dat ik Jan het project regelmatig gedurende een tijdje zag hinderen*
that I Jan the project frequently for some time saw hamper

This means that, apparently, the trace of the raised verb in (22) cannot precede the PP, as in (23a). The interpretation of (22) forces the analysis in (23b), with the trace following both the adverbial and the PP. As noted, in case of two prehead modifiers, precedence determines c-command. Hence the unambiguous scopal relation in (22).

(23) a. **dat ik [Jan het project regelmatig t_i gedurende een tijdje]*
 [zag hinderen_i]
 b. *dat ik [Jan het project regelmatig gedurende een tijdje t_i]*
 [zag hinderen_i]

So, while there is no adjacency condition as such on rightward incorporation, as shown above, adjacency is required in (23). This follows directly from the parsing strategies we have assumed. The fact that the embedded verb cannot raise across one of its own dependents is a consequence of the fact that the LB specification of such a dependent would have to be destroyed, as discussed in section 4. The parser will assign an LB specification to the XPs it encounters before encountering the verb cluster, as in (24a).

- (24) a. $XP_{LB} XP_{LB} XP_{LB}$
 b. $XP_{LB} XP_{LB} XP_{LB} t_i V V_i$
 b'. $XP_{LB} XP_{LB} t_i XP_{LB} V V_i$

When the parser encounters the second verb and hypothesizes it has been moved, a trace must be inserted. However, none of the LB specifications may be altered. Suppose the trace is inserted after all XPs, as in (24b). In that case, these XPs can be analyzed as dependents of the moved verb without any problem. Since they precede the trace, they can be on left branches of its projection. Suppose now the trace is inserted before one of the XPs, as in (24b'). In that case, this XP can only be analyzed as being a dependent of the matrix verb. If it were a dependent of the embedded verb, it would have to be on a right branch, as it follows the trace. Hence the possibility of (17d) versus the impossibility of (23a).

6. CONCLUSION

We have argued that base-generated structures are in principle symmetrical. With respect to movement, however, an asymmetry can be observed. Head movement to the right is impossible when the head crosses its own dependents, while for head movement to the left (for instance in cases of V2) this is no problem. We have argued that this asymmetry is a consequence of parsing strategies, which in turn crucially depend on the symmetry of syntactic structures.

NOTES

* We would like to thank Arthur Dirksen, Frank Drijkoningen, Hans van de Koot, Jan Odijk, Tanya Reinhart, Eddy Ruys, Fred Weerman and the audiences at the 1996 'Linguistics in the Netherlands' colloquium, the University of Cologne and WECOL 1996.

1. With double object constructions, a cross-linguistic anti-mirror image effect can be observed; see Neeleman & Weerman 1997 for discussion.
2. We abstract away from the possibility that two adjacent XPs form a constituent. By 'XP' we refer here to those maximal projections that are specifiers, complements or adjuncts of the head of the construction, ignoring possible XP-internal complexity.
3. String-vacuous rightward head movement (namely V-to-C movement in the strictly head-final languages Japanese and Korean) is argued for by Whitman (1991). Whitman shows that certain properties shared by V2 languages and Japanese/Korean can be explained by assuming V-to-C in both.
4. Again, there is no ban on raising the embedded verb across dependents of the matrix verb. However, this is hard to illustrate in the case of V-to-V raising. The example in (ia) seems to be an instance of raising across a matrix PP. Unfortunately, due to clause union effects that result from V-to-V raising, PPs that are apparently generated in the embedded clause can get a matrix construal as well, as shown by (ib). This means that the evidence provided by (ia) is inconclusive.

- (i) a. dat Jan [Marie de samba t_i] van een afstandje [zag dansen_i]
that Jan Marie the samba from a distance saw dance
- b. dat Jan [Marie van een afstandje de samba t_i] [zag dansen_i]
that Jan Marie from a distance the samba saw dance

REFERENCES

- Berwick, Robert, and Amy Weinberg. 1985. Deterministic parsing: a modern view. In *Proceedings of NELS 18*, eds. Stephen Berman, Jae-Woong Choe and Joyce McDonough, 15-33. Amherst: GLSA.
- Evers, Arnold. 1975. *The transformational cycle in Dutch and German*. Doctoral dissertation, Utrecht University.
- Frazier, Lyn. 1987. Syntactic processing: evidence from Dutch. *Natural Language and Linguistic Theory* 5:519-559.
- Frazier, Lyn. 1993. Processing Dutch sentence structures. *Journal of Psycholinguistic Research* 22:83-108.
- Frazier, Lyn, and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: eye movements and the analysis

- of structurally ambiguous sentences. *Cognitive Psychology* 14:178-210.
- Gorrell, Paul. 1995. *Syntax and Parsing*. Cambridge: Cambridge University Press.
- Greenberg, Joseph. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language* (2nd edition), ed. Joseph Greenberg, 73-113. Cambridge, Mass.: MIT Press.
- Hoeksema, Jack. 1991. Complex predicates and liberation in Dutch and English. *Linguistics and Philosophy* 14:661-710.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Koster, Jan. 1974. Het werkwoord als spiegelcentrum. *Spektator* 3:601-618.
- Marcus, Mitchell, Donald Hindle and Margaret Fleck. 1983. D-theory: talking about talking about trees. In *Proceedings of the 21st annual meeting of the Association for Computational Linguistics*, 129-136. Cambridge, Mass.: MIT.
- Neeleman, Ad. 1994. *Complex predicates*. Doctoral dissertation, Utrecht University.
- Neeleman, Ad and Fred Weerman. 1992. Case theory and the diachrony of complex predicates in Dutch. *Folia Linguistica Historica* 13:189-217.
- Neeleman, Ad and Fred Weerman. 1997. *Case and arguments in a flexible syntax*. Ms. Utrecht University/OTS.
- Reuland, Eric. 1990. Head movement and the relation between morphology and syntax. In *Yearbook of morphology 1990*, eds. Geert Booij and Jaap van Marle, 129-161. Dordrecht: Foris.
- Rochemont, Michael. 1992. Bounding rightward A'-dependencies. In *Island constraints*, eds. Helen Goodluck and Michael Rochemont, 373-397. Dordrecht: Kluwer.
- Sturm, Arie and Johan Kerstens. 1978. Over verklaren in de taalkunde. In *Nieuwe tegenstellingen op Nederlands taalgebied*, eds. Peter van Putte and Henk Verkuyl, 130-172. Utrecht: Instituut de Vooys.
- Whitman, John. 1991. String vacuous V to COMP. *GLOW Newsletter* 26:62-63.

**Mostly Predictable:
Cyclicity and the distribution of schwa in Itelmen¹**

Jonathan David Bobaljik, Harvard and McGill

Abstract. The Itelmen language shows many instances of regular schwa/zero alternations. Even though the language permits quite extensive consonant clusters, I argue in this paper that the alternating schwas (and perhaps all instances of schwa) are epenthetic; schwa is inserted to break up a disfavoured consonant cluster. The rule which inserts schwa must apply cyclically in the verbal system, but non-cyclically in the nominal system. Apparent examples of cyclic rule application are problematic for non-derivational, constraint-based approaches to phonology, such as many versions of Optimality Theory (OT). Thus, much recent work has been devoted to reanalyzing purported examples of cyclicity from an O.T. perspective. While it may be possible to devise an account of the Itelmen data in terms of parallel constraint evaluation, current O.T. approaches are insufficient; in particular, the best candidate for an explanation of the Noun/Verb differences (Base Identity) makes exactly the wrong predictions for Itelmen.

Introduction. Itelmen (a.k.a Kamchadal) is today spoken natively by fewer than 100 people, living on the Northwest coast of the Kamchatka peninsula of Russia. Genetically, it is likely related to the geographically proximate Chukchi, Koryak, Kerek and Alutor, though it shows striking differences from these languages both in terms of the lexicon and at a typological level. Of interest for the present paper is the distribution of schwa in Itelmen. That there are many instances of schwa-zero alternations (1), suggests that these occurrences of schwa may involve epenthesis.²

¹ I am most deeply indebted to the speakers of Itelmen who shared their time, hospitality, and knowledge with me, especially (but not only): T.N. Bragin†, N.I. Chatkina, A.D. Ivashova, L.E. Pravdoshchina, S.E. Prichin, A.E. Shamuraeva†, E.E. Silina, I.I. Yaganov† and N.S. Yaganova (Northern dialect), V.P. Krasnoyarev†, V.V. Prichin†, N.V. Prichina, M. P. Slabodchikova, E.P. and V.D. Zaporotskaja, G.D. and N.Z. Zaporotskij, and D.N. Zhirkov (Southern dialects). The work reported on here represents a part of a larger, on-going project. Many people have commented on the larger project; for comments specific to this paper I wish to thank Susi Wurmbrand, Glyne Piggott, Ingvar Löfstedt, Michael Kenstowicz, Morris Halle, Edward Flemming, Noam Chomsky and other members of the audiences at WECOL (UC Santa Cruz), McGill and MIT. The omission of many of their suggestions clearly detracts from the present work, and I hope to incorporate these into the larger project. For funding, I acknowledge the support of a National Council for Soviet and East European Research (grant to D. Koester) and of the Milton Fund of Harvard University. I would also like to thank Wilson Gray for tracking down a copy of Moll 1960—an extremely valuable and hard to obtain article on Itelmen dialects. Errors of fact, interpretation and anything else are mine alone.

² Examples are taken primarily from notes and recordings from three trips to Kamchatka (1993-94, spring and summer 1996), supplemented with examples from Volodin 1976 and Volodin and Khaloimova 1988. When there are relevant dialect differences, forms marked (N) or (S) are from the Northern and Southern dialects, respectively. Examples are representative of the material which has been carefully transcribed and/or rechecked with speakers. Though I believe the generalizations to be true of the language generally, the reader is cautioned that not all of the recordings have been carefully transcribed.

With the following exceptions, examples are given in IPA: s.z are apical, post-alveolar fricatives (the underdot is omitted for convenience); sequences written as a glottal stop and nasal (e.g., ?n) or ?l correspond to a single, glottalized segment and not a series of two segments; words

- | | | | | | | |
|-----|----|------------|----------------|-----|-----------|----------------|
| (1) | a. | ʔxəm | sable | a'. | ʔxm-ɛʔn | sable-PL |
| | b. | spəl | wind | b'. | spl-ank | wind-LOC |
| | c. | *ʔxəz-xʔal | road-ABL | c'. | *ʔxz-enk | road-LOC |
| | d. | il:-əz-in | drink-PRES-3SG | d'. | ʔxu-z:-in | stand-PRES-3SG |

The descriptive goal of this paper is to argue that the instances of schwa in the forms in the left column of (1) are inserted to break up otherwise-ill-formed consonant clusters. At first blush, such an approach might appear difficult to maintain since Itelmen regularly permits extensive consonant clusters, word-internally and at word edge (2).

- | | | | | |
|-----|------------|---------------|----------|-----------------------|
| (2) | čkrəč | 'spoon' | tʰsčŋin | 'you are carrying it' |
| | kʔzukeʔn | 'they were' | mskčɛʔn | 'I will make them' |
| | sitʔxpk'ɛʔ | 'with embers' | k'ənsʔxč | 'boil it!' |

I will show nevertheless (section 1) that these alternating schwas occur in discrete, identifiable contexts and are thus predictable. Moreover, I will show that most instances of non-alternating schwa in the language occur in the same environments as the alternating schwas and could thus be predicted by the same epenthesis rule. Though the epenthesis rule to be motivated is simple enough, in order to derive the correct forms, the rule must apply cyclically in verbs and non-cyclically in nouns. As cyclicity effects of this type are potentially problematic for non-derivational approaches to phonology (such as many current versions of O.T.), section 2 will be devoted to a discussion of the theoretical implications of the assumptions required in section 1. Finally, a brief appendix considers three classes of apparent exceptions, identifying the specific environments defining these classes and arguing that they are not true counter-examples to generalizations made here.

1. ON THE DISTRIBUTION OF SCHWA.

As a point of departure, we delineate a certain class of consonants which I will refer to as R (e.g., resonants) in what follows. These consonants are given in (3a). The remaining consonants I will group perhaps inaccurately as K (-obstruents) in (3b) for comparison. Note that there are four additional consonants which never appear in positions where their behaviour relative to schwa epenthesis is testable.³

preceded by a superscript * are pronounced round throughout—this rounding is morphological and cannot be tied to any particular segment or segments.

The following abbreviations are used in this paper: ADjective, ADVerb, ABLative, DIMinutive, PEJORative, LOCative, ASPECT, ASP2=second aspect, FUTure, PREsent, INFinitive, IRRealis mood, PRT=participle, NEGative suffix, SUBject, OBJect, CLitic, SG=singular, PL=plural.

Discrete morphemes are separated by hyphens; the dot in the glosses separates distinct features expressed on a single portmanteau morpheme. Where a gloss has the form, e.g., 2pl>3pl it indicates a portmanteau agreement marker, here second person plural subject acting on third person plural object.

³ I use the terms "resonant" and "obstruent" somewhat imprecisely here, and avoid them in the general discussion, referring instead just to "R". In standard terms, the voiced apical fricative /z/ is [-sonorant] and hence not a resonant; the contrast between /l/ ∈ R versus /ʎ/ ∉ R is also potentially curious. Note, though, that the same classification of segments relative to similar

- (3) a. R = { m, n, ŋ, l, r, z }
 b. K = { p, p', t, t', k, k', q, q', č, č', φ, x, xʷ, χ, s, t }
 c. untestable = { j, β, γʷ, ? }

1.1 Nouns Taking first the alternating schwas in the nominal system (1a-c), schwa always appears in the environment ...C_RC or ...C_R#, that is, following a consonant and immediately preceding a stem-final R which in turn is followed by a consonant-initial suffix or word-end. When the R-final noun stem is followed by a vowel-initial suffix (as in the prime examples in (1a'-c')), then no schwa surfaces. This situation clearly suggests that schwa is epenthesized for reasons of syllabification: although the language permits extensive consonant clusters, consonants from R must be immediately adjacent to a vowel.⁴ The epenthesis rule is schematized in (4).⁵

- (4) $\emptyset \rightarrow \text{ə} / \left\{ \begin{smallmatrix} \text{C} \\ \# \end{smallmatrix} \right\} \text{---} \text{R} \left\{ \begin{smallmatrix} \text{C} \\ \# \end{smallmatrix} \right\}$

Confirmation for the approach taken here comes from the distribution of non-alternating, word-internal schwa in the language. Laying aside certain exceptions discussed below, word-internal schwa in Itelmen occurs in the environment proscribed by (4); conversely, there are no cases of R occurring between two consonants, unsupported by schwa (i.e., *CRC).⁶ Examples of non-alternating schwas are given in (5a): schwa occurs only when the R can not be the coda onset of a full vowel; when R can syllabify in this manner, no schwa occurs (5b).

- (5) a. əmŋəl qetit-kənkin isxəmt-lax
 story freeze-NEG.PRT dull-ADJ
- b. č'amzanl-aʔn *č'aməzanəlaʔn
 person-PL

properties (i.e., syllabification) is motivated for certain Salish languages (see Matthewson 1994 on St'át'imcets, among others).

⁴ One argument that this is schwa epenthesis and not syllabic resonants (and /z/) comes from the spread of palatalization. In Itelmen, /č/ and /ʃ/ are always palatalized. In general, /n/ and /l/ contrast with palatalized /nʲ/ and /lʲ/ in all positions, but these segments are always palatalized before the inherently palatalized segments. This spreading of palatalization does not cross schwa. For example, the word for 'fish' in the Northern dialect is [nən'č]—only the second /n/ is palatalized by /č/ even though there is no general prohibition against /nʲ/ in onset, cf. [nʲen'eqečχ] 'child-DIM.'

⁵ There are certain complications to do with word-initial exceptions. See the appendix, below.

⁶ St'át'imcets, a Salish language in which extensive consonant clusters are also attested, likewise prohibits R from cluster-medial position, requiring epenthetic schwa. For description and analysis, see Matthewson 1994. While Itelmen is strikingly similar to the Salish languages in a number of ways, there are notable differences concerning, e.g., the interaction of epenthesis and stress, and the possibility of obstruent only words (see, e.g., Bagemihl 1991).

1.2 Verbs The verbal domain is slightly more complicated than the nominal domain for two reasons. To begin with, there are no schwa/zero alternations in verb roots. In verb roots, occurrences of final R not preceded by a full vowel are always preceded by schwa; unlike noun roots, the initial segment of the following suffix is irrelevant. This is illustrated in (6); note that the root *sp(ə)l-* 'wind' can inflect either as a noun (1b) or verb.⁷

(6)	t-zəl-čen 1SG-give-1>3SG 'I gave it (to him).'	xan-zəl-um=nen 3.IRR-give-1SG.OBJ=CL 'She might give me (to him)'
	spəl-qzu-in wind-ASP-3SG 'It was windy'	spəl:-in⁸ wind-3SG 'It is windy'

The left-hand examples in (6) pattern with the nouns in (1a-d): schwa surfaces before a stem-final R followed by a suffix-initial consonant. However, as illustrated in the right-hand column, this schwa does not disappear when the suffix is vowel-initial, as would be predicted by (4). In other words, schwa in the environment C__R] in verb roots never alternates with zero, even though it does so in cognate noun roots (compare the noun and verb for 'wind'). Importantly, there is no minimal requirement on the shape or size of verb stems which could motivate schwa epenthesis here; verb roots with no vowel are common:

(7)	m-sk-čeʔn 1SG.IRR-make-1>3PL 'I will make them'	k-ʔ-qzu-kn-eʔn PRT-be-ASP-PRT-PL 'they were'
-----	--	---

While there are no schwa/zero alternations in verb roots in Itelmen, there are alternations in the verbal inflectional morphology. The present tense suffix has four predictable allomorphs: [-əz-], [-əs-], [-z-] and [-s-]. The choice among these involves two alternations: voiced vs. voiceless and schwa vs. zero. The voiced/voiceless alternation is a straightforward case of regressive devoicing—if the segment immediately following the tense inflection begins with a (voiceless) consonant, then the voiceless allomorphs surface, while if the following suffix begins with a vowel, then the voiced alternant is used—(8a) vs. (8b):⁹

⁷ Note that very few roots show a dual life of this sort. Most roots are uniquely nominal or verbal.

⁸ Gemination is generally predictable: R → R: / V'___ (i.e., post-tonic, intervocallic R's lengthen). For complications which gemination may pose, see Appendix, section A.3.

⁹ Exceptional in this regard is the behaviour of /ʔ/. When a root or affix terminating in /ʔ/ immediately precedes the present tense marker, the /ʔ/ and /z/ somehow coalesce into a single surface segment [s]. Schwa is never epenthesized before this segment, and the segment remains voiceless even if, on the surface, it comes to be between two vowels. The infinitive, past and present of the verb meaning 'come, arrive' are given in (i):

(i) k'ot-kas INF k'ot-in (PAST)-3SG k'o-s-in -PRES-3SG

Note that [ʔs] clusters are generally permitted in the language, as in q-sunʔ-sx (2.IRR-live-2PL.SUBJ).

- (8) a. t'-nu-s-kičēn b. nu-z:-in
 1SG-eat-PRES-1SG eat-PRES-3SG
 'I am eating.' 'He/she is eating.'

While this voicing alternation is sensitive to the environment to the right of the present tense marker, the schwa-zero alternation is sensitive exclusively to the environment to the *left* of the tense marker. If the verb stem (including pre-tense inflectional markers, if any) terminates in a vowel, then there is no schwa (9a-c), but if the verb stem ends in any consonant—except /h/, see fn. 9—including any member of R, then schwa is obligatorily present (9d-f).

- (9) a. t-tχzu-s-kičēn b. ʔeru-z-in c. ʔqu-z-in
 1SG-stand-PRES-1SG gripe-PRES-3SG be-ASP-PRES-3SG
 'I am standing' 'she gripes' 'she is'
- d. t'-il:-əs-kičēn e. il:-əz-in f. spəl:-əz-in
 1SG-drink-PRES-1SG drink-PRES-3SG windy-PRES-3SG
 'I am drinking' 'he drinks' 'it is windy'

Ignoring for a moment the devoicing of /z/, the occurrence of schwa in (9d) is predicted by (4): (underlying) R sandwiched between consonants requires epenthetic schwa for reasons of syllabification. By the same token, though, schwa in (9e-f) is unexpected; since the following segment is a vowel, /z/ should be able to syllabify as an onset, as in comparable examples from the nominal system (1a'-c'). Example (9f) is particularly striking when contrasted with (1b'): in the nominal form there is no schwa epenthesis, and in the verbal form, there are two schwas.

The correct forms are derived on the assumption that the rule of epenthesis (4) applies cyclically in verbs and non-cyclically in nouns, illustrated below.¹⁰

¹⁰ An obvious question concerns the behaviour of stems in derivations which involve category changing, such as nominalizations of verbs. Unfortunately, the language, as far as I can tell, conspires against us. One test case would require a vowel-initial nominalizer added to the R-final verb stem (i.e., if the nominalizing suffix is consonant-initial, the environment for epenthesis is met on the surface for nouns as well as verbs). I have found no such morphemes. Another case would involve a vowel-initial verbal (derivational) suffix immediately following the verb root, followed in turn by a nominalizing suffix. Again, I have been unable to construct relevant cases, due to independent properties of the language's morpheme inventory.

(10)	cyclic	non-cyclic
a. spl 'wind' + -ank (LOC)	[[spl] ank]	spl-ank
Cycle 1	[spəl]	
Cycle 2	[[spəl] ank]	
OUTPUT	* spəlank	splank
b. spl + PRES + 3SG	[[[spəl] z] in]	spl-z-in
Cycle 1	[spəl]	
Cycle 2	[[spəl] əz]	
Cycle 3	[[[spəl] əz] in]	
OUTPUT	spələzin	*spəlzín, *spləzín

On the first cycle of a cyclic derivation, the root alone is evaluated. Since following suffixes are not visible until subsequent cycles, a root-final R will trigger epenthesis if preceded by a consonant. This is the incorrect result for nouns, but is the correct result for verbs, deriving the fact that the schwas in ...CR] verb stems such as those in (6) do not alternate.

For the verb (10b), the present tense suffix /z/ is added on a subsequent cycle (the third line in (10b)). When this morpheme is added to a consonant-final stem, the environment for epenthesis is met—/z/ is a member of R, it is preceded by a consonant, and is followed by nothing (i.e., since the next suffix is added on the next cycle).

As shown in (10), the cyclic derivation makes the wrong prediction for nouns, predicting no alternation of schwa with zero before stem-final R. A single, non-cyclic application of (4) correctly predicts the alternations seen in the noun system. However, the inverse situation obtains for verbs. The non-cyclic derivation would, for instance, incorrectly predict the schwas in (6) to alternate with zero. Moreover, the cyclic derivation not only correctly predicts the alternations in the present tense marker, but it also correctly predicts the lack of alternation in verb roots.

An additional piece of evidence, alluded to above, also points to the cyclic nature of Itelmen verbal derivations. Recall from (8) that the devoicing of the present tense morpheme is essentially predictable from the nature of the following segment. Note moreover that schwa is epenthésized before the present tense marker (after consonant-final stems) regardless of whether or not the present tense marker is devoiced by a following consonant (see, e.g., (9d)). What is important about this observation is that the voiceless apical fricative, /s/, is not a member of R—that is, devoicing should bleed the application of epenthesis. This can be seen word-internally: while /z/ is not permitted between consonants, /s/ does not trigger epenthesis and occurs freely in clusters:

(11)	k-sk-kna-ʔn	əŋqs-q-aʔ-iʔn	k-ənsxt-iʔn
	PRT-do/make-PRT-PL	hurt-ASP2-FUT-3PL	PRT-bear-PRT.PL
	'they appeared'	'they will hurt'	'they bore them'

Crucially, regressive devoicing must apply *after* schwa epenthesis. This follows straightforwardly on the cyclic derivation: the environment for schwa

epenthesis is met at the point when the present tense suffix is added; the environment for devoicing does not arise until the subsequent cycle.

To summarize, three generalizations emerge from consideration of the Itelmen data considered here.

- (12) a. Most instances of schwa are in the environment: $\{C\}_{\#} _ R \{C\}_{\#}$.
 b. There are no surface sequences $\{C\}_{\#} R \{C\}_{\#}$.
 c. All instances of schwa not described by (12a) are in verbs and are in the environment C_R-V (where "-" indicates a morpheme boundary)

For each of these generalizations, there is a corresponding class of exceptions. These are considered in turn in the Appendix; there, I will show that these do not constitute counter-examples to the characterization of the facts as presented here, but do motivate some refinement. In brief, one class of apparent exceptions involves schwa which is inserted as a last resort in words with only obstruent consonants—a sort of minimal word effect. The other two classes of exceptions involve the left periphery of the word: while certain root-initial Rs fail to trigger epenthesis, we see that root-initial consonants behave specially in other respects as well; finally, a conflict arises with certain cases of stressed schwa before geminate, intervocalic R—the schwa epenthesis is predictable if the geminate is underlying, but the gemination would be predictable if schwa is underlying, an uncomfortable state of affairs for the theory, but not a crucial counter-example. Again, there are clear considerations involved for each of the classes of apparent exceptions to (12), and discussion of these is postponed until the appendix.

Throughout this section, I have argued that the observed generalizations motivate a simple analysis of the distribution of schwa in Itelmen, whereby schwa is always epenthetized for reasons of syllabification, as given in (4). A crucial assumption is that syllabification and epenthesis work cyclically in verbs and non-cyclically in nouns (recall the derivations in (10)). Further motivation for the cyclic nature of verbal derivation comes from the interaction of epenthesis and devoicing, particularly the failure of devoicing to bleed epenthesis in the present tense marker (9d).

Examples of apparent cyclic rule application of this sort pose a significant problem for non-serial, constraint based approaches to phonology such as many current versions of O.T. For instance, the environment for schwa epenthesis in verb stems and in the present tense marker obtains only at an intermediate level of representation, and is not met in either the underlying form or an output string which would lack epenthesis, yet it is exactly these intermediate levels which certain versions of O.T. explicitly deny. These facts must be captured in some different manner. In the next section, I will show that approaches to this problem in the current O.T. literature do not suffice for the Itelmen cases, lending support to proponents of cyclic derivations in phonology.

2. THEORETICAL IMPLICATIONS

One thread of recent work within the O.T. paradigm has focussed on attempts to capture apparent examples of cyclic effects in a non-serial, single-level model. Within this body of research, two families of approaches have emerged: one which appeals to alignment constraints requiring morphological and phonological

constituents to match up (see, among others, McCarthy & Prince 1993, Kenstowicz 1994, and references therein); and one which develops forms of output-output correspondence constraints (see Benua 1995, Burzio 1995, Kenstowicz 1995). In this section, I will evaluate to what degree selected proposals from these families are able to handle the Itelmen syllabification data. We will see that, where the predictions of current models are clear, they are wrong for Itelmen. For some proposals, it is not clear that no account could in principle be available, and I will only discuss what hurdles an account would have to overcome to be descriptively adequate for the syllabification data presented above. Since the focus of this section is to consider alternatives to cyclic models, I will have nothing to say about serial models, whether derivational or representational (such as multi-stratal O.T.).

2.1 Base Identity Of the various alternatives to cyclicity in the mono-stratal O.T. models, the only approach I am aware of which seeks to explain and predict differences between nouns and verbs is the Base Identity approach proposed by Kenstowicz 1995. Hence, we begin the discussion with this model. Kenstowicz's formulation of the relevant constraint is given in (13).

- (13) **Base Identity:** Given an input structure [X Y] output candidates are evaluated for how well they match [X] and [Y] if the latter occur as independent words. (Kenstowicz 1995:8)

The function of this constraint is to promote some candidate on the basis of its similarity to an independently occurring word in the language, even though some other output candidate would otherwise be more favourable given the constraints operative in the language. This is perhaps best explicated by means of Kenstowicz's examples. Consider first the various forms of a noun (14a) and verb (14b) from a conservative dialect of Korean, given in (14)

- | | | | | | | |
|------|----|----------|------------|----|-----------|---------------------|
| (14) | a. | /kaps/ | 'price' | b. | /Eps/ | 'not have' |
| | | kap | citation | | _____ | no uninflected form |
| | | kaps-i | nominative | | Eps-Ess-E | past-informal |
| | | kap-k'wa | N-'and' | | Ep-t'a | non-past-formal |

The noun and verb each have a final consonant cluster in their underlying representations; this cluster surfaces before a vowel-initial suffix (third line of each example). Before a consonant-initial suffix (fourth line), or word-finally (second line), the cluster simplifies—the final consonant is left unparsed.

Consider now the corresponding forms from the speech of younger speakers, given in (15).

- | | | | | | | |
|------|----|----------|------------|----|-----------|---------------------|
| (15) | a. | /kaps/ | 'price' | b. | /Eps/ | 'not have' |
| | | kap | citation | | _____ | no uninflected form |
| | | kap-i | nominative | | Eps-Ess-E | past-informal |
| | | kap-k'wa | N-'and' | | Ep-t'a | non-past-formal |

The difference between the two dialects surfaces in the third line—the vowel-initial suffixes. For the more conservative speakers, the underlying consonant clusters in both nouns and verbs surface in exactly this environment. However, for the younger speakers, the underlying cluster surfaces only in the verbs; in the nominal system, the underlying cluster fails to surface regardless of

phonological environment.¹¹ Thus, there is an apparent difference in syllabification between nouns and verbs. Base Identity gives a straightforward account of this difference: even though the stem-final /-s/ in the noun in (15a) is in principle syllabifiable before the vowel-initial nominative suffix (third line), Base Identity prefers an output without this /-s/ since such an output (i.e., *kapi*) corresponds more closely to the citation form *kap*, an independently occurring lexical item.¹² The crux of the account comes from the observation that verb stems, unlike noun stems, cannot occur in their bare form. Since there is no corresponding, uninflected form of the verb, Base Identity is irrelevant in the verbal system in Korean. In this way, differences of behaviour for syllabification between the nominal and verbal system are explained in terms of Base Identity since in only one case (the nouns) does the base correspond to an independent lexical item.

Since the noun stem may stand alone, Base Identity predicts that syllabification throughout the nominal paradigm should be uniform; the verbal paradigm should show sensitivity to the first segment of the following suffix, since there is no independent word corresponding to the bare verb stem. In Itelmen, as in Korean, noun stems may surface in their uninflected form (for Itelmen, subjects and direct objects surface thus), while verb stems may not surface without inflectional affixes. Thus, Base Identity should have the same effect in Itelmen as it does in Korean, predicting invariance in the nominal system—the stem form should not change regardless of the following suffix—and variability in the verbal system. This is the opposite of what we have seen above to be the case. This is illustrated in the following tableaux.

(16) Noun
'wind'+ LOC

Base Identity	* $\left\{ \begin{matrix} C \\ \# \end{matrix} \right\} R \left\{ \begin{matrix} C \\ \# \end{matrix} \right\}$	Fill
spəl 'wind'		

spl-ank <--correct form	*!		
spəl-ank WRONG ! ✗			*

¹¹ Kenstowicz (1995) does not discuss the motivation for including the cluster in the underlying representation in (15), as opposed to an alternative whereby the younger speakers have reanalysed the underlying forms as having no cluster. As the present purpose is merely to exemplify the mechanics of Base Identity, this question is academic in this context.

¹² The relevant constraints and their ranking are: *Complex >> Base Identity >> Parse-C, where *Complex bars complex onsets or codas, and Parse-C says that a consonant in the input form must show up in the output. I have omitted tableaux for reasons of space; see Kenstowicz 1995.

(17) Verb 'be windy'+ PRES + 3SG	Base Identity n/a	*{C}R{C}	Fill
-------------------------------------	----------------------	----------	------

spl-z-in		*!	
spəl-z-in	WRONG ! ✗		*
spl-əz-in	WRONG ! ✗		*
spəl-əz-in	<-- correct		**!

In these tableaux, the rule in (4) is translated as two constraints, one against unsyllabifiable R, the other, *Fill*, disfavouring epenthesis. Obviously, *Fill* must be the lower ranked of the two or there would be no epenthesis regardless of environment. Similarly, in order to have any effect whatsoever, Base Identity must be ranked above *Fill*; its ranking with respect to the *CRC constraint is irrelevant. As can be seen clearly from the tableaux, this approach predicts that schwa epenthesis which is not motivated by surface environment should only be motivated to conform with an independent stem, exactly the wrong prediction for both verbs and nouns.¹³

Kenstowicz's (1995) Base Identity is the only approach which attempts to derive verb vs. noun contrasts of this sort. Unfortunately, we have seen that the approach is not able to provide an account of the Itelmen data. Recent literature includes other output-output correspondence approaches, including *Uniform Exponence* (Kenstowicz 1995) or *Anti-Allomorphy* (Burzio 1995). Informally, these constraints require "minimiz[ing] the differences in the realizations of a lexical item (morpheme, stem, affix, word)" (Kenstowicz 1995:20). Issues of complexity aside (Itelmen verbs inflect for 2 moods, 3 tenses, 2 conjugation classes, and agreement in person and number with subject, direct and indirect object, in addition to a rich array of productive derivational morphology) such approaches have nothing to say about the systematic difference between nouns and verbs. More conclusively, *Anti-allomorphy* would have to be relativized to the verb root by stipulation. Recall that noun roots do show allomorphy, and in the verb system there are four allomorphs of the present tense marker: [-əz-], [-əs-], [-z-] and [-s-]. Given the cyclicity effects, separate constraint rankings would be necessary in order to drive the differences in environments for epenthesis between noun stems and the present tense marker. Moreover, relativizing an anti-allomorphy or a similar constraint to apply only to the verb root would reduce to a restatement of the descriptive observation which cyclicity explains, and is thus unsatisfactory.

2.2 Generalized Alignment. The second family of approaches for dealing with apparently cyclic phenomena in a monostratal O.T. model captures the effects of morpheme boundaries by requiring, e.g., that some edge of a morphological

¹³ I have used the verb and noun from the root *spl* 'wind' for purposes of exposition. Most roots in Itelmen are solely verbal or solely nominal and cannot inflect across categories in this way. Hence, an appeal to Base Identity between the verb stem and the bare noun will work accidentally in the case of (17), but cannot work as a general property of the system since most verbs do not have a cognate bare noun.

constituent coincide with the edge of a phonological constituent, a family of constraints originating in McCarthy & Prince (1993). One could imagine an appeal to Generalized Alignment which would handle the cases which the output-output correspondence theories failed to handle, in particular, the case of multiple epenthesis of schwa as in (17). Saying nothing about the nouns, some constraint (referred to below as GA) requires that the right edge of verbal roots and of the present tense marker—but mysteriously of no other suffix—coincides with the right edge of a syllable. Moreover, though complex codas are allowed, some Coda constraint prohibits CR codas. If both of these constraints are ranked above the constraint disfavouring epenthesis, then the correct candidate in (17) is predicted.

(18) Verb
'be windy'+ PRES + 3SG

CODA	GA	FILL
* CR	+v],R , σ,R	

a. spl-z-in	*!	**	
b. spəl-z-in		*! (-z-)	*
c. spl-əz-in		*! (stem)	*
d. spəl-əz-in	☞		**!

In this system, violations of the coda constraint can be avoided if, e.g., in (18b) the /z/ syllabifies as an onset to the agreement suffix, or in (18c) if the final /l/ of the verb root syllabifies as an onset to the syllable headed by an epenthetic schwa. However, in each case a violation of the GA constraint would be incurred. Only by epenthesis more than once could both the Coda and GA violations be avoided. Technically, this provides an account of the forms which Base Identity could not account for, but such an approach raises other questions. One implication is that the morpheme-final consonants, even intervocalically, are not syllabified as onsets; in (18d) or structurally parallel examples with full vowels, the syllables are VC and lack onsets.¹⁴ There also remains the question of why it is only these suffixes which must be listed in the formulation of GA. More importantly, this approach may run into problems with vowel-final verb stems, as illustrated in (19):

¹⁴ Though it is often assumed that onsets are preferred over codas universally, evidence that at least some languages are best analysed as having VC syllables is given in Breen & Pensalfini, to appear.

- (19) Verb Vowel-Final stem
'stand'+ PRES + 3SG

CODA * CR	GA +v],R ,σ,R	FILL
--------------	------------------	------

a.	$\begin{array}{c} \sigma \quad \sigma \\ \wedge \quad \wedge \\ \text{t}\chi\text{z}\text{u} \quad \text{z} \quad \text{in} \end{array}$		*! (-z-)	
b.	$\begin{array}{c} \sigma \quad \sigma \\ \wedge \quad \wedge \\ \text{t}\chi\text{z}\text{u} \quad \text{z} \quad \text{in} \end{array}$		*! (stem)	
c.	tχzu-əz-in	Wrong CR		*

The problem which this appears to pose is that when the present tense marker is intervocallic, syllabification of /z/ as an onset (19a) incurs a GA violation since the present tense suffix is not aligned with the right edge of a syllable. Avoiding this by syllabifying the present tense marker as a coda (19b) likewise incurs a GA violation since the verb stem is not aligned to the right edge of a syllable. The winning candidate in such cases would appear to be the one with schwa epenthesis, contrary to fact.¹⁵

A final problem for the GA approach is the failure of regressive devoicing to bleed schwa epenthesis (see the discussion above (11)). On the surface, /s/ should not trigger epenthesis—indeed, /-s/ in root-final position or in suffix initial position never forces schwa epenthesis. Both facts are illustrated in (20):

- (20) qomsqzuβumsx < /q-oms-qzu-βum-sx/
2.IRR-abandon-ASP-1SG.OBJ-2PL.SUBJ
'Leave me (here)'

3. Conclusion. The descriptive goal of this paper has been to show that the distribution of schwa in Itelmen is predictable from simple considerations of syllabification, considerations which are familiar from other languages. The caveat 'mostly' in the title refers to a certain case of potential exceptions, discussed in the appendix below. In particular, the Itelmen epenthesis data—and its relationship to a rule of devoicing—points to the role of the cycle in providing a concise account of Itelmen phonological processes.

Since apparent cyclic behaviour of this sort is potentially problematic for mono-staratal approaches to phonology, such as many versions of O.T., the second half of the paper was devoted to an exploration of the implications of the Itelmen data for current theories. Two families of recent non-serial alternatives to O.T. were considered: output-output correspondences, and Generalized Alignment. For the first group, the one analysis which attempts to derive noun versus verb asymmetries of exactly the sort examined here (Base Identity, Kenstowicz 1995), makes the wrong predictions for both the verbs and the nouns. While I am aware of no particular analysis of similar data invoking Generalized Alignment, in section

¹⁵ Edward Flemming suggests that a highly ranked constraint against vowel-schwa sequences, for instance, may allow one or both of (19a,b) to win.

2.2 I sketched what such an analysis might look like. Like the Output-Output Correspondence approaches, GA appears to face significant problems which do not arise on a cyclic approach. As always, this type of argument runs the risk of setting up an implausible account and then pointing out the flaws in it, failing to understand the true potential of a given theory. Even with this in mind, I hope to have shown that there may be a more general problem with the GA account.

On an account which embraces a cycle (including serial versions of O.T.), it remains unexplained why nominal derivations are non-cyclic and verbal derivations cyclic. Nevertheless, from this one assumption, the facts presented above follow straightforwardly: schwa in verb roots never alternates with zero, alternating schwa in nouns is sensitive to both preceding and following segments while the alternating schwa in the present tense suffix is sensitive only to the preceding segment, and regressive devoicing of the present tense suffix fails to bleed schwa epenthesis. The GA approach likewise has no account of the difference between nouns and verbs; however, there is no equivalent assumption to cyclicity, and this clustering of properties can no longer be seen to constitute a natural class with a unified explanation. Each must be derived by different combinations of constraints.

It is undoubtedly the case that some account in a mono-stratal approach will mechanically derive the facts presented here; the challenge for such an approach is to capture at the same time the observation that the cycle expresses a generalization, unifying a range of phenomena as the result of a single aspect of the phonological component. To the extent the descriptive generalizations are true, one would hope that they fall out of the theory.

APPENDIX: THE EXCEPTIONS

For the sake of completeness, I will briefly consider here the apparent exceptions to (12) and argue that these are not counter-examples for the analysis presented here, but rather form clearly delineable classes of principled exceptions, which for the most part find plausible explanation from general considerations. At the very least, the fact that the exceptions all concern the left edge of the word or the first syllable suggests that they are not accidental by any means.

A.1 CəC The first class of exceptions concern (12a), and involve occurrences of schwa between two consonants not from R. Some examples are given in (21).

(21)	a.	ʷqəsx	dog	a'.	ʷqsχ-aj	dog-PEJOR
	b.	čkəp	fungus	b'.	čkp-əʔn	fungus-PL
	c.	čkpəč	spoon	c'.	čkpč-əʔn	spoon-PL
	d.	čət	gun, bow	d'.	čt-əʔn	gun-PL
	e.	kəpkəp	tooth	e'.	kp-əʔn	tooth-PL

In each of these examples, the schwa alternates with zero in suffixed forms, e.g., with the plural or pejorative. More importantly, in the forms on the left, there are no vowels other than schwa, nor are there any consonants from R in the word. And, with the exception of reduplicated forms such as (21e), each form involves only one schwa. The data thus points to a "last resort epenthesis"—a requirement that all words have at least one syllable. In this way, the forms in (21) would count as a principled exception to the approach advocated here and are not problematic.

A.2 #RC Another class of exceptions consists of Rs which are not adjacent to vowels (and thus cannot be syllabified) but nevertheless fail to trigger epenthesis, violating (12b). Some of these are listed in (22)

(22)	lq-lax	cold-ADJ	mč'emč	rowan berry
	lq-aq	cold-ADV	msetenəŋ	stick used for digging
	zlatumx (N)	sibling	ŋrepe-z-in	sing-PRES-3SG

In all cases which I have found of this sort, the unsyllabified R is the initial consonant of a root. While no explanation springs immediately to mind, the fact that all of the cases involve a root-initial R suggests that the direction to an explanation lies with other edge-phenomena such as extrametricality. Of possible relevance here is a historical split between the Northern (Sedanka) and Southern (Khairiuzovo) dialects of Western Itelmen, first noted by Moll 1960. While both dialects maintain a voiced-voiceless contrast for fricatives in root-initial position (23b-c), there are certain mismatches where a voiced fricative in the Northern dialects correspond to a voiceless fricative in the Southern dialects (23d-e).

(23)	(gloss)	North	South
a.	thimble	βolβol	βolβol
b.	give	zəl-	zəl-
c.	which	sag	sag
d.	seal	βitβit	φitβit
e.	live	zunt-	sunt-

These idiosyncratic voicing mismatches between the dialects occur only in the initial position of certain roots, the same position that the idiosyncratic extrasyllabic Rs occupy.¹⁶ Hence, the exceptions form a discrete class, with some independent plausibility in the context of the language, even if no account is forthcoming.

A.3 əR:V The final class of exceptions involve schwa before R which, it would appear, could syllabify as the onset to a full vowel, as in (24).

(24)	kəm:a (S)	1sg pronoun	səl:atumx (S)	sibling
	əz:o-s	put-INF	qəl:al	snow
	əl:e-kas	get-INF	əz:-ank	outside-LOC

One point of regularity about these cases is that they all occur in the first syllable of the word. As Itelmen has initial stress, these schwas all receive word stress. Moreover, in all of these cases, the R following the schwa is geminate. In

¹⁶ That the correct characterization of the position in question is root-initial, as opposed, e.g., to word-initial, can be seen when inflectional prefixes are added, thus:

- (i) k'-ŋrep-qzu-knen PRT-sing-ASP-PRT.3SG
- (ii) t-zun-s-kičen (N) 1SG-live-PRES-1SG
- (ii') t-sun-s-kičen (S) --same--

this fact, however, lies the root of a paradox. One could propose that the geminates are underlyingly two segments, in which case the appearance of schwa in (24) would be predicted (i.e., by (4)). On the other hand, gemination in Itelmen is otherwise fully predictable: post-tonic, intervocalic R is always geminate, and this environment is the only position in which geminates are attested. Cases such as (24) would be no exception to the rule of gemination if the schwa is underlying. And thus, the paradox. If it were not for cases like (24), both gemination and the distribution of schwa would be entirely predictable, but in cases like (24), it appears that either gemination or epenthesis, but not both, is predictable.

To summarize, the generalizations in (12) are not without exceptions. However, what I have shown in this section is that the exceptions fall into three classes, each clearly delimitable and thus understandable in terms of other aspects of the grammar. The first class is easily explained by appeal to a minimality constraint on words motivating a last resort epenthesis. The two remaining classes of exceptions both involve the left periphery of the word: extrasyllabic root-initial consonants, and exceptionally marked geminates or first syllable schwas. While an appealing account is not readily forthcoming, the fact that we can so clearly delineate the classes of exceptions suggests that an account can be found.

REFERENCES

- Bagemihl, Bruce. 1991. Syllable structure in Bella Coola. *Linguistic Inquiry* 22.4, 589-646.
- Benua, Laura. 1995. Identity Effects in Morphological Truncation. To appear in J. Beckman, S. Urbanczyk & L. Walsh, eds., *UMOP 18: Papers in optimality theory*.
- Breen, Gavin & Rob Pensalfini. To appear. Arrernte: a language with no syllable onsets. *Linguistic Inquiry*.
- Burzio, Luigi. 1995. Surface constraints versus underlying representation. Ms. The Johns Hopkins University.
- Kenstowicz, Michael. 1994. Cyclic versus non-cyclic constraint application. In Andrew Carnie, Heidi Harley and Tony Bures, eds., *MIT Working Papers in Linguistics, vol. 21, Papers on phonology and morphology*. Pp. 11-42.
- . 1995. Base Identity and Uniform Exponence: alternatives to cyclicity. To appear in J. Durand & B. Laks, eds., *Current trends in phonology: Models and methods*, CNRS Paris X and University of Salford, University of Salford Publ.
- McCarthy, John & Alan Prince. 1993. Generalized Alignment. In Geert Booij and Jap van Marle, eds., *Yearbook of morphology 1993*. Kluwer, Dordrecht.
- Matthewson, Lisa. 1994. Syllable structure in St'at'imcets. To appear in *Proceedings of CLA, 1994*, Toronto Working Papers in Linguistics.
- Moll, T.A. 1960. Очерк фонетики и морфологии седанкинского диалекта ительменского языка. Ученые записки Ленинградского государственного педагогического института им. А. И. Герцена, кафедра языков народов крайнего севера. [An essay on the phonetics and morphology of the Sedanka dialect of Itelmen. *Scientific notes*. Department of the languages of the peoples of the far North, A. I. Herzen State Pedagogical Institute, Leningrad.]
- Volodin, Aleksandr Pavlovich. 1976. Ительменский язык [The Itelmen language]. Leningrad: Nauka.
- & Klavdia Nikolaevna Khalojmova. 1989. Словарь ительменско-русский и русско-ительменский [Itelmen-Russian and Russian-Itelmen Dictionary] Leningrad: Prosveshchenie.

**Relations between core and periphery in
American Sign Language phonology
Diane Brentari
University of California—Davis**

1. Introduction

Most work on the phonology of sign languages has focused primarily on what is often called the core component of the lexicon, which is that part whose forms: 1) are monomorphemic, and 2) show no synchronic traces of borrowed elements from English through the fingerspelling system. This eliminates from consideration all initialized forms, forms that have evolved from reduced fingerspelled forms, as well as forms that are the result of the concatenation of classifier forms and movement roots, yet it is in these components that expansion of the lexicon primarily takes place. The first study that investigates the nativization operations of fingerspelled forms is found in Battison (1978), and this account is the basis for much of the subsequent work on this subject. The paper here proposes that the native lexicon is heterogeneous, but in systematic ways, and that the non-native portions of the lexicon weaken constraints that hold in the native part.

Itô and Mester (1995a, 1995b) have proposed a model of the Japanese lexicon that is based on principles that are directly relevant to an analysis of the ASL lexicon. In Japanese, the Yamoto forms constitute the native sub-component; the Sino-Japanese, Foreign, and Mimetic sub-components are peripheral. Itô and Mester (1995a, 1995b) argue that many of the constraints that hold for the native sub-component cease to hold or are weakened in systematic ways in the peripheral sub-components. They also predict that: 1) the sub-components of the lexicon do not behave as non-overlapping entities within the grammar, but rather that principles of the core are weakened in peripheral sub-components; 2) peripheral sub-components do not add or strengthen a constraint; and 3) that the sub-components should be identifiable by differences in segmental inventories and exploitation of constraints. These predictions are empirically supported by forms in ASL.

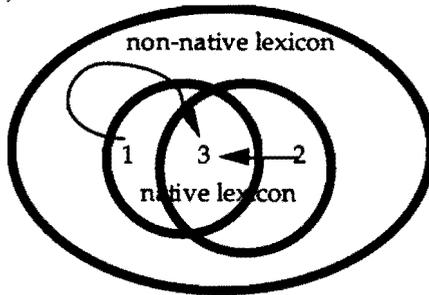
The first distinction I wish to make is between the native and non-native lexicon of ASL, and show the structural ways that these two broad classes can be defined (1). The native lexicon includes: 1) the fingerspelled alphabet; 2) the polymorphemic predicates, often called classifier predicates, that are the syntactic clauses and prosodic words; and 3) the 'frozen forms,' which I will refer to as the "core" lexicon. To be a member of the non-native component, the form has some evidence of fingerspelling.¹

(1) Native Lexicon in ASL

1. *the fingerspelled alphabet*: a set of handshapes, a few of which are also specified for orientation, orientation change, or movement (i.e., ASL names for the orthographic letters (e.g., 'K', 'P' (orientation), 'J' (orientation change), 'Z' (movement))).

¹There are other types of forms that undergo nativization, such as forms from other sign languages or non-manual elements, but those types of borrowings are not addressed in this paper.

2. *Classifier predicates*: polymorphemic system made up of bound roots and a variety of types of affixes (e.g., 2 stooped, upright beings, side by side, facing forward, move forward carefully, from 'a' to 'b').
3. *Core lexicon*: a repository of forms whose source can be from the classifier predicate system directly, or from the fingerspelled alphabet after a form has conformed to a set of 'nativization' constraints (e.g., BREAD, TO-FLY, AIRPLANE).



The fingerspelled alphabet (part 1), is a set of names for the English alphabet, consisting primarily of handshapes, a few of which are also specified for orientation (e.g., 'G', 'Q', 'U', 'H', 'K', 'P',) or for movement ('J' has an orientation change, and 'Z' has a tracing movement). I would draw the analogy between fingerspelled alphabet in ASL as a part of the language and that of the different words that the orthographic letter 'y' has in various spoken languages ([wai] in English, [igrEk] in French, [ipsilōw] in Portuguese, [ipsolon] in Italian). The fingerspelled alphabet is the set of ASL names for the orthographic letters, and in some sign languages, this is the limited role that they play, used as frequently (or infrequently) as we spell out words in English. In many sign languages (e.g., French, Dutch, German, Danish, Italian, Japanese, Chinese), the mechanism for borrowing words from the dominant surrounding language is not a system like the fingerspelled letters, but some other means, and ASL signers are thought to overuse fingerspelling. The polymorphemic component (part 2) is made up of bound roots and a variety of types of affixes that can be put together to form classifier predicates (Supalla, 1982). These structures remain a productive part of the native lexicon and are identified by their general lack of person inflection and derivational morphology and by their spatial agreement properties (Padden, in press). These include the ability to allow derivational and inflectional morphology appropriate for the semantics of the stem of person, number, and aspect to affix the forms. Two examples of such stem and affix forms in the core, one from Padden and Perlmutter (1987) and one from Padden (1988), are: ACT, which has a derived nominal form ACTING (activity nominalization through affixation of a feature [trill] to the stem); the verb GIVE, which has a number of forms that exhibit person and/or agreement—1-GIVE-3/3-GIVE-1 (showing person agreement by a change in the beginning and end points of the sign), or GIVE [pl. object] (showing affixation of an arc movement to the stem).

The core lexicon is fed by both the fingerspelling system, through a complex set of operations of nativization to be discussed in the rest of this paper, and directly by the classifier predicate system. For example, the 'AIRPLANE' handshape is a productive bound classifier morpheme that can be put together in the

polymorphemic part of the lexicon in forms, such as 'to-dock-in-outer-space'; the stem is also the basis for the verb TO-FLY and the derived noun AIRPLANE, which, because of this morphological property, can be placed in the core lexicon. Classifier predicates are verb phrases morpho-syntactically, and these forms allow neither person/number inflection or derivational morphology, but instead encode primarily spatial reference.

2. Differences among components of the native lexicon

At this point I should introduce some terminological distinctions concerning handshape and movement (2). *Selected fingers* (Mandel, 1981, Sandler, 1989; Brentari, 1990b, in press) is that specification capturing those fingers that are able to move or contact the body during the course of executing a sign.² *Joints* (van der Hulst, 1995, Brentari, in press) is that specification capturing the joints of a handshape that may flex. The positions are open (no joints specified), bent (non-base joints specified) curved (base and non-base joints specified), flat (base joints specified), and closed (whole hand closed in a fist). *Aperture* specifies whether the joints are open or closed. With respect to movement, there is a distinction between *path movements*, executed by the elbow and shoulder joints, and *local movements*, articulated by the wrist and knuckle joints.

(2) Terms defining handshape and movement distinctions

handshape:

Selected fingers: specification capturing those fingers that are able to move or contact the body during the course of executing a sign.

Joints: specification that captures the joints of a handshape that may be flexed. The specifications are open (no joints), bent (non-base joints), curved (base and non-base joints), flat (base joints), and closed (whole hand closed in a fist).

Aperture: specification for the joints as open or closed.

movement:

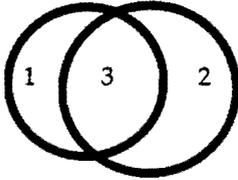
Path movements: input movements executed at the elbow and/or shoulder.

Local movements: input movements of the wrist and/or knuckles.

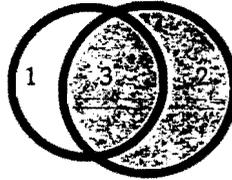
The native lexicon of ASL is not unified, but consists of these three subcomponents. We have evidence for their autonomy if we examine the handshape inventories of them (3). There is considerable correspondence among these three native subcomponents, as seen in (3a), but I would like to call attention to the non-overlap between the set of handshapes in part 1 (i.e., the fingerspelled alphabet) and that of parts 2 and 3 (i.e., the classifier predicates and core forms).

² This definition is not comprehensive, since these criteria do not apply to every sign. For example, many signs have no contact nor handshape change in the sign (e.g., HURT, PREACH, BAPTIZE, etc.). Also, in the signs GERMANY and CIGAR, contact occurs at a place other than the so-called selected fingers.

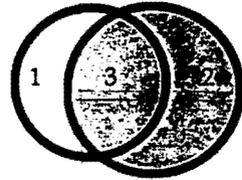
(3) Handshape asymmetries in the native lexicon

a. correspondence
1=2=3

'B'(open)(curved)
'H'(open)
'l'(open)(bent)
'A'
'S'

b. non-correspondence
2,3≠1 (additons)

'D'
'E'
'M'
'K/P' contrast
'G/Q' contrast
'V/K' contrast
'U/H' contrast

c. non-correspondence
2,3≠1 (omissions)

'B'(bent)(flat)
'H'(bent)(curved)(flat)
'l'(flat)(curved)
'horns'
'7'
'8'

Examining the fingerspelled alphabet for handshapes that it lacks (3c), 'B' (bent)(flat), 'H' (bent)(curved)(flat), 'l'(flat)(curved), and 'horns' appear in the classifier system and in core forms, but not in the fingerspelled alphabet. The fingerspelled alphabet also has some handshapes and handshape contrasts that the other two parts of the native lexicon does not have (3b). 'D', 'E', 'M' are in the fingerspelled forms, but not elsewhere in the native lexicon. Also, 'K' and 'P', 'U' and 'H', and 'Q' and 'G' are contrastive by virtue of a specified orientation in these fingerspelled letters, and 'K' and 'V' are contrastive because of a feature [stacked], but these pairs of handshapes are not contrastive in the classifier predicate system or in the core lexicon; they are allophonic. To see a chart of the fingerspelled alphabet and a more complete set of examples of this non-overlap, see the appendix.

3. Constraints in the native lexicon.

The relation of movements to syllables is summarized in (4). The number of sequential movements is equal to the number of syllables, and a movement can consist of one or more local or path movements as defined in (2); the set of allowable handshape changes will be discussed below. All ASL words must have at least one movement, hence words are at least one syllable long (Brentari 1990a, b, c, in press, Perlmutter 1992). There are a set of four constraints that hold in the native lexicon (5), which are violated in particular ways in the non-native lexicon.

(4) Relations among Prosodic words, Syllables, and Movements

Movements may be a local or path movement.

The number of syllables in a word equals the number of sequential movements in a sign.

All lexical items must contain at least one movement.

(5) Constraints that hold in the native lexicon

2-MVT: (2 movements/word maximum)

There are at most two movements (i.e., syllables) per prosodic word.

2-HS: (2 handshapes/word)

Prosodic words contain at most two handshapes.

SF: (restrictions on changes in selected finger groups)

a. one selected fingers group per prosodic word.

b. Hand-internal movements involve only selected fingers.

MAX-AP: (Maximize aperture change)

Handshape changes that occupy syllable peaks maximize aperture change.

I want to elaborate on each of the constraints in (5), explaining just how they came to be formulated as they are here. The constraint **2-MVT** captures the fact that all words in ASL are maximally two sequential movements (Perlmutter 1992, Brentari 1994). The status of the polymorphemic classifier predicates as prosodic words or prosodic phrases is questionable, and because classifier predicates may contain more than two movements, I will assume that they are not prosodic words, and say no more about them. The **2-HS** constraint captures the fact that prosodic words, regardless of whether they are mono- or disyllabic, contain a maximum of two different handshapes. There are disyllabic forms with three output handshapes (e.g., DESTROY, REMOVE, CERTIFY), but there are only two different handshapes in these forms; the first and last handshapes are the same. Many constraints on morpheme-internal handshape sequences (Sandler, 1989), syllable-internal sequences (Brentari, 1990b; Perlmutter, 1992) and word-internal sequences (Friedman, 1977; Mandel 1981) have been proposed. The **SF** constraint places two restrictions on the kind of handshapes these two handshapes must be in the core lexicon in the simplest possible way. **SF(a)** says that core words should contain the same selected fingers, and **SF(b)** says that only the selected fingers can move during the course of executing a core word. The **MAX-AP** constraint focuses on the requirements for a well-formed syllable with respect to handshape aperture. It captures the fact that, *ceteris paribus*, hand-internal, local movements contain one open and one closed handshape whenever possible (Brentari, 1994). The **MAX-AP** constraint is a re-formulation of an earlier constraint in Brentari (1990b), called the **PERIPHERALITY CONSTRAINT**, which said that prosodic words can contain a maximum of one bent, curved, or flat handshape. Instead of stating the joint specifications in the constraint itself, **MAX-AP** works with an inventory of handshape pairs based on their joint specifications, argued for in Brentari (in press) (6).

(6) Open and closed handshapes

a. with all fingers selected

open-closed	RUN-OUT-OF
curved-closed	BITE
bent-open	WANT
flat-open	INFORM
closed-open	SEND

b. in the fingerspelled alphabet

closed (bent, flat, closed)
'M', 'N', 'O', 'S', 'A', 'T', 'X', 'E'
'D' (nonselected), 'F' (selected)

open (open, curved)
'H', 'I', 'J', 'K', 'L', 'Q', 'P', 'R',
'U', 'Y', 'Z', 'B', 'C', 'G'
'D' (selected), 'F' (nonselected)

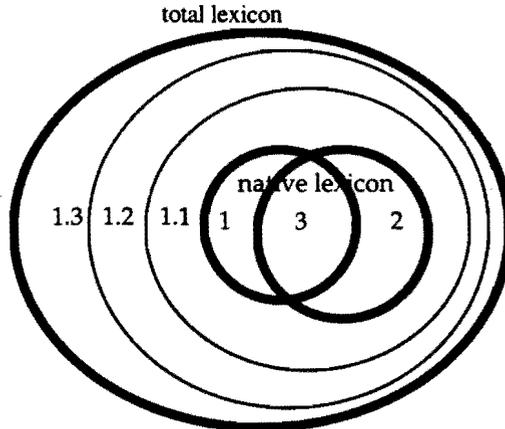
The three constraints **MAX-AP**, **2-HS**, **SF** are a reformulation of two constraints in Perlmutter (1992)—a 2-type constraint, limiting the number of different ‘types’ of handshapes in a lexeme to two (i.e., sets of selected fingers), and a handshape contour constraint, limiting the number of ‘tokens’ (i.e., joint specifications or aperture settings) in a syllable to two. By separating the number of handshapes from the type of handshapes present in native lexical items using **MAX-AP** and **SF**, I would claim that violations in the non-native lexicon are rendered more transparent. Furthermore, both of Perlmutter’s constraints miss the point that no matter whether the two handshapes are tokens or types, they conform to sequences of open and closed handshapes.

4. The non-native lexicon

This section will focus on signs which contain some fingerspelled elements, and I provide an analysis for dividing them up into four groups (7): 1) those that are a part of the native lexicon; and those that are part of the non-native lexicon: 1.1) those that violate the selected fingers constraint **SF**; 1.2) those that violate **SF**, **MAX-AP**, and **2-HS**; 1.3) forms that violate **SF**, **2-HS**, **MAX-AP** and **2-MOV**.

(7) The non-native lexicon in ASL

- 1.1 Some arbitrary name signs, abbreviated name signs, initialized signs, ≤ 2 -letter loans. These are forms that violate **SF**.
- 1.2 Partially assimilated loan signs, ≤ 3 -letter loans. These forms violate **SF**, **MAX-AP**, and **2-HS**.
- 1.3 Commonly fingerspelled words, sign+fingerspelled compounds. These forms violate **SF**, **2-HS**, **MAX-AP**, and **2-MVT**.



The types of signs in each of the non-native strata are those proposed by Padden (in press) and are somewhat heterogeneous; however, I wish to divide them according to structural behavior alone whenever possible. As Battison does in his 1978 analysis, I use one of the non-native classes of signs which has members in all strata of the non-native lexicon as a basis for its stratification—*loan signs*. I

then use the structural differences that have emerged to locate other types of non-native forms in the established strata. I define loan signs as signs which: 1) contain fingerspelled letters; 2) that occur in a restricted place of articulation, in the so-called 'neutral' space in front of the signer rather than with respect to a specified place on the body; and 3) that have a restricted set of movements. They may contain more than two handshapes. For example, BUT (B-T) WOULD (W-D), WHAT (W-T), JOB (J-B) contain two handshapes, but EASY (E-S-Y) and EARLY (E-R-L-Y) contain more than two. In addition to loan signs, there are other types of non-native vocabulary that the analysis here will explain. *Arbitrary name signs* are generated by combining the first initial of the person's first (or first and family) English name in a set of restricted places of articulation in combination with a restricted set of movements—a trilled movement, one or two straight movements, or a movement that is the enhanced form of the local movement. In *initialized signs* and *abbreviated signs*, the handshape of the initial letter of the English word is combined with the movement and place of articulation of a native form. As Padden (in press) describes them, initialized signs are often, but not always, members of a semantic field. Two examples are COMPUTATION, with a 'V/K' handshape, and the related initialized signs STATISTICS, ALGEBRA, CALCULUS, GEOMETRY, and TRIGONOMETRY; another example is SCIENCE, with an 'A' handshape, and the related initialized signs BIOLOGY, CHEMISTRY, and EXPERIMENT. This semantic generalization does not apply to abbreviated forms (e.g., 'feedback' (F-B), 'videotape' (V-T), 'withdraw' (W-D)). One important difference between the set of loan signs and the sets of arbitrary name signs, abbreviated signs, and initialized signs is that all of the fingerspelled letters of the English word are a part of the input to loan signs, while this is not the case in arbitrary name signs, abbreviated signs, and initialized signs. Finally there are forms in ASL that are simply fingerspelled in their full form, either alone (e.g., unfamiliar place names and some company names) or in combination with a sign, as in the *sign+fingerspelled compounds* L-E-G+WORK 'legwork', SUN+B-U-R-N 'sunburn'). See the appendix for a more complete list of more of each type of these signs.

The movements in these forms are systematically different, though I do not treat them formally here. The movements and places of articulation in *abbreviated* and *initialized* signs are taken from a core form. For example, the abbreviated form 'social work' (S-W) takes its movement and place from HELP; the initialized form BIOLOGY takes its movement and place from SCIENCE. *Loan signs* (e.g., J-B, 'job', B-T 'but', S-R-E 'sure'), which employ no movements from core form, may employ epenthetic straight movements to a point of contact or movements involving phonetic enhancement of the transitions between fingerspelled letters (Stevens, Keyser and Kawasaki 1986, Stevens and Keyser 1989) by adding an appropriate path movement to the local movement articulated by the sequence of handshapes. See Brentari (in press) for a discussion of translation statements which allow a movement normally articulated as a local movement to appear as a path movement in the output. *Arbitrary name signs* may employ straight movements to point of contact, or trilled movements for signs with no contact. Except for the constraint 2-MOV, I leave the formulation of the movement constraints for future research.

Now let us turn to the stratification of the lexicon using stable loan words as a base. Most ASL words that contain fingerspelled letters are not part of the core component of the lexicon, but here I analyze some forms containing remnants of the fingerspelled alphabet as core forms because: 1) they contain exclusively

fingerspelled letters in input; 2) they conform to all the constraints in (5); and 3) the handshapes of the output conform to the set of handshapes in parts 2 and 3 of the core lexicon. This happens by means of 'handshape merger', whereby two or more handshapes have been combined into one; these cases were first discussed in Battison (1978). By understanding how the constraints operate in core forms, we will be able to apply them consistently to other types of forms.

The constraint tableau for BREAD is given in (10). At this point, I need to present two more types of constraints, both of which are common to analyses in Optimality Theory—ALIGNMENT and FAITHFULNESS. ALIGNMENT is that set of constraints that matches up prosodic categories, such as syllables, with morphological categories, such as stems. In this case ALIGN(L) matches the first letter of stems with the first handshape of the word, and ALIGN(R) matches the final letter of the English word with the final handshape of the word. FAITHFULNESS constraints do their best to guarantee that the shape of the output matches the input to the greatest extent possible. They militate against deletions from the input form—MAX constraints—or against epenthesis in the output form—DEP constraints. The only FAITHFULNESS constraint I use is called MAX-HS, which requires that all handshapes of the input must also be present in the output. ALIGN and MAX-HS are given in (8) and (9). To make the points I am about to make clear, I have used the term FAITH instead of the label MAX-HS in the following tableaux.

(8) ALIGN constraints

- A. ALIGN(L): initial handshape of stem with left edge of stem
- B. ALIGN(R): final handshape of word with right edge of word

(9) MAX-HS: all handshapes of the input must be present in the output.

The tableaux in (10)-(13) show how the core and non-native forms behave with respect to the proposed constraints. The only constraint that must be crucially ranked in the core form BREAD is FAITH. All of the other constraints are unviolated in the output form, '8' [open][closed]³; the other forms are less harmonic in the following ways. The fully fingerspelled form BREAD (the first candidate) has two MAX-AP violations because 'B' and 'R' are both 'open' and 'E' and 'A' are both 'closed'. This form contains five, not two, handshapes, thereby incurring three violations of 2-HS. There are three violations of SF(a) because while 'B' and 'E' contain the same selected fingers, 'R', 'A', and 'D' cause three changes in selected finger groups. SF(b) is violated only once, and ALIGN(L) and ALIGN(R) are unviolated, as is FAITH. Candidates 2 and 3 with unparsed first and last letters incur one less 2-HS and SF(a) violation, but they violate ALIGN(R) and ALIGN(L) respectively. Deleting two of the middle

³ Two explanatory remarks about the well-formed candidate 7 are in order. One is that the movement realized in the output form is rapidly repeated. The second is that '8' is a conventionalized, shorthand way of notating a handshape with the middle finger selected and nonselected fingers open. The handshape change [open] -[closed] with this set of selected fingers indicates an output form which changes from having all of the fingers extended (just like 'B') to one where the middle finger is flattened, and the index finger and the ring + pinkie finger are extended independently. The extended index finger approximates the 'D'.

handshapes (candidate 4) eliminates the violations of **MAX-AP** and **2-MOV**, but the violations of **SF** and **2-HS** remain. The forms with two handshapes (candidates 5 and 6) are quite well-formed, 'B-D' incurs only one violation of **SF(a)**, but the actual output fairs even better, having no violations of the proposed constraints except for **FAITH**, with a whopping five violations. Notice, too, that the only form with no violations of **FAITH** is the fully fingerspelled form.

(10) Constraint tableau for **BREAD**(8[o][cl]) (nativized form)

In loan signs with two handshapes, (such as 'J-B' (11)) **FAITH** is no longer dead last. The output has a violation of **SF(a)** and **SF(b)**. It is equivalent in structure and number of violations incurred with the 'B-D' form of **BREAD**—candidate 6—except that 'J-B' has not rid itself of the fingerspelled letters. The output contains one violation of **FAITH**. The tableau for **EASY**, a non-native form in stratum 1.2, shows **FAITH** moving up further in the constraint hierarchy. Violations of **2-HS** and **SF** are what define this stratum. In stratum 1.3, **FAITH** is ranked about all constraints except **ALIGN(L)** and **ALIGN(R)**. In this stratum any violation of **FAITH** will be fatal. In (14), we can trace the degree of faithfulness to the input of loan signs with respect to the constraints of the core lexicon. It is important to re-iterate that forms can be stable members of these strata.

/BREAD/	2-MOV	ALIGN-L	MAX-AP	2-HS	ALIGN-R	SF.a	SF.b	FAITH
BREAD	**		**	***		***	*	
BREA<D>	*		**	**	*	**	*	*
READ	*	*	**	**		**	*	*
B<RE>AD				*		**		**
BA					*	*		***
BD						*		***
+8[o] [cl]								*****

(11) Constraint tableau for 'job' (J-B) (non-native, part 1.1; loan sign)

/JOB/	2-MOV	ALIGN-L	MAX-AP	2-HS	ALIGN-R	FAITH	SF.a	SF.b
JOB			*				*	*
+J-B						*	*	*

(12) Constraint tableau for 'easy' (E-S-Y) (non-native, part 1.2; loan sign)

EASY	2-MOV	ALIGN-L	ALIGN-R	FAITH	MAX-AP	2-HS	SF.a	SF.b
EASY	*				**	**	*	
E-Y				**			*	*
+E-S-Y				*	*	*	*	*

(13) Constraint tableau for 'stock' in S-T-O-C-K MARKET (non-native, Part 1.3; sign+fingerspelled word)

STOCK	ALIGN-L	ALIGN-R	FAITH	2-MOV	MAX-AP	2-HS	SF.a	SF.b
+STOCK				**	**	**	*	
S-K			***		**		*	

(14) The ranking of FAITHFULNESS in native and non-native components of the ASL lexicon

Native (1.0)	Non-native (1.1)	Non-native (1.2)	Non-native (1.3)
2-MOV	2-MOV	2-MOV	ALIGN(L)
ALIGN(L)	ALIGN(L)	ALIGN(L)	ALIGN(R)
MAX-AP	MAX-AP	ALIGN(R)	FAITH
2-HS	2-HS	FAITH	MAX-AP
ALIGN(R)	ALIGN(R)	2-HS	2-MOV
SF(a)	FAITH	MAX-AP	2-HS
SF(b)	SF(a)	SF(a)	SF(a)
FAITH	SF(b)	SF(b)	SF(b)

5. Combinations in other native forms

Thus far in the loan signs discussed, there are no combinations of forms from two sub-components of the native lexicon; there are only fingerspelled letters. Initialized and abbreviated signs combine movements of core forms with one or two fingerspelled letters, respectively. Padden (in press) discusses some restrictions on these combinations. For example, she notes that if a handshape in a core form retains its status as a classifier, a fingerspelled letter may not be substituted for it. Signs violating these combinatoric restrictions are judged ungrammatical by native ASL signers, and many such forms occur in manually coded English systems. I would argue that another combinatoric restriction of this type is seen in the ASL name sign system. Name signs are generated to refer to members of the Deaf community and arise via a mechanism for developing "classifier name signs" and "arbitrary name signs" (Supalla 1992). Classifier name signs are generated by combining a bound movement root and a classifier handshape in a place of articulation (often depicting a salient aspect of the person's habits or personality); arbitrary name signs are generated by combining the first (and sometimes last) initial of the person's English name, given the restrictions on movement and place of articulation described earlier. Native ASL signers reject as ungrammatical name signs often invented by hearing signers, which combine bound movement roots of the classifier name signs (part 3 of the native lexicon) with handshapes of arbitrary name signs (part of the native lexicon).

According to the constraints argued for in this paper, combinations of a 'movement from a core form+one fingerspelled letter' that crucially do not violate the combinatoric restrictions sketched above are potentially members of the core lexicon. *Mutandis mutandi*, forms with two fingerspelled handshapes would be members of stratum 1.1 of the non-native lexicon.. In (15) we see the tableau for an initialized form with two handshapes—W-S 'workshop'. The candidate set helps us see a crucial ranking between ALIGN(L) and ALIGN(R). Because of the 2-HS constraint, all of these forms allow for two empty handshape slots in the

input, in addition to the path movement from the core form. What we see here is that a two-handshape form incurring one violation of ALIGN(R) is preferred over a form incurring one violation of ALIGN(L). The preferred form chooses the two leftmost handshapes of the two English stems.

(15) 'workshop'+GROUP (path) (non-native, Part 1.1; abbreviated sign)

'workshop'+GROUP	ALIGN-L	ALIGN-R	2-HS	MAX-AP	SF.a
+W-S+GROUP		*			*
W-P+GROUP	*				*

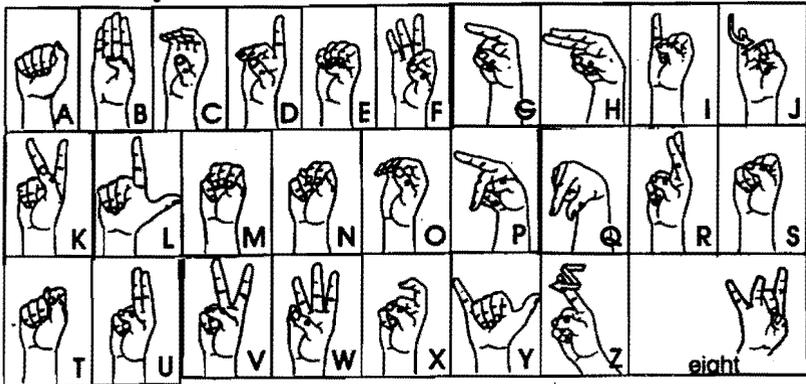
6. Conclusion

This model is meant to be language-specific, and it is with respect to operations of lexicalization and mechanisms for borrowing that I would expect sign languages to differ. There is growing evidence that this is the case. For example, many European sign languages use non-manual components, often called 'Mundbilder' as a mechanism for borrowing from the surrounding spoken language, and in the Far East, sign languages can borrow from the surrounding language by means of the Chinese characters which are ascribed visually similar handshapes in the sign language in question.

One reason lexical innovation is important as grammatical evidence in the case of a sign language is that ASL is relatively young, and another is that the cultural profile of the Deaf community in the United States has changed dramatically in a short time. The number of Deaf individuals in the professions of medicine, law, veterinary science, dentistry, business, and research (in the sciences and in the humanities) has increased dramatically in recent years, creating a need for specialized vocabulary. I hypothesize that the mechanisms for lexical innovations will employ existing well-formedness principles of the language.

Appendix

I. The manual alphabet



II. Examples of overlap and non-overlap in handshape inventories within the native lexicon

a. 1=2=3:

(1)fingerspelled	alphabet	(2)classifier predicates	(3)core forms
H(open)	H, U, V	'flat, slender object'	TRAIN
S	S	'spherical object'	DENTIST
A	A	OPEN-NEWSPAPER	SECRET
B(open)	B	'flat surface'	BROKE
B(curved-o)	C	'vertical round object'	COMPLAIN
B(curved-cl)	O	'round, thick, flat object'	NOTHING
l(open)	Z	'upright being'	BITTER
l(bent)	X	'hunched, upright being'	PERPLEXED
F(1-flat.cl)	F	'small, round object'	CAT
L	L	'rectangular, flat object'	RUN
R	R	'twisted, long strings'	CIGAR
Y	Y	'fat, by-legs'	WHY
V	V	'by legs'	SEE
G	G	'slender, flat object', or 'thick, flat object'	BEFUDDLED/ CLUELESS

b. 2,3≠1 (additions)

(1)fingerspelled	alphabet	(2)classifier predicates	(3)core forms
D	D	***	***
E	E	***	***
M	M	***	***
N	N	***	***
K	K	allophone of V	allophone of V
P	P	allophone of V	allophone of V
Q	Q	allophone of G	allophone of G

c. 2,3≠1 (omissions)

(1)fingerspelled	alphabet	(2)classifier predicates	(3)core forms
'horns'	***	'horns'	MOCK
B(bent)	***	'claws'	WANT
H(bent)	***	'bent, flat, slender object'	WHISTLE
l(curved)	***	'curved, slender object'	RED-FACED
H(curved)	***	'curved, flat, slender obj.'	SNAKE
l(flat)	***	'handle, tiny object'	BIRD
H(flat)	***	'flat, slender object'	NO
B(flat)	***	'flat, wide object'	MALE
'8'	***	'absence' (???)	HATE

III. Loan signs

J-B	'job'	W-T	'what' (Exclamation)
E-S-Y	'easy'	B-T	'but'
Y-S	'yes' (emphatic)	S-R-E	'sure'
E-R-L-Y	'early'		

IV. Abbreviated signs

V-T	'videotape'	F-B	'feedback'
P-J	'project'	B-G	'background'
S-W	'social work'	W-S	'workshop'
S-C	'senior citizen'	B-T	'board of trustees'

V. Initialized signs

Based on GROUP (C handshape)		Based on PERSON (B handshape)	
TEAM	GROUP	SUBJECT	HUMAN
ASSOCIATION	SOCIETY	CLIENT	PERSON
DEPARTMENT	FAMILY		
Based on COMPUTATION (K handshape)			
TRIGONOMETRY	GEOMETRY		
ALGEBRA	STATISTICS		
CALCULUS			

VI. Sign+fingerspelled compounds (from Padden, in press)

B-E-L-L BOY	'bellboy'
L-E-G WORK	'legwork'
P-R-O-O-F READ	'proofread'
F-O-O-T WORK	'footwork'
S-T-O-C-K MARKET	'stock market'
DEAD E-N-D	'dead-end'
SUN B-U-R-N	'sunburn'
SOAP B-O-X	'soapbox'
EYE T-O-O-T-H	'eyetooth'
CHEAP S-K-A-T-E	'cheapskate'

References

- Battison, Robbin. (1978). *Lexical borrowing in American Sign Language*. Silver Spring, MD: Linstock Press.
- Brentari, Diane. (1990a). *Theoretical foundations of American Sign Language phonology*. University of Chicago dissertation. Published 1993, University of Chicago Occasional Papers in Linguistics.
- Brentari, Diane. (1990b). Licensing in ASL handshape. In C. Lucas (Ed.), *Sign language research: Theoretical issues*, 57-68. Washington, DC: Gallaudet University Press.
- Brentari, Diane. (1990c). Underspecification in American Sign Language phonology. In K. Hall, J-P. Koenig, M. Meacham, S. Reiman, and L.A. Sutton (Eds.), *Proceedings from the 16th annual meeting of the Berkeley Linguistics Society*, 46-56.
- Brentari, Diane. (1994). Prosodic Constraints in American Sign Language. In S. Gahl., A. Dalbey, and C. Johnson (Eds.) *Proceedings from the 20th annual meeting of the Berkeley Linguistics Society*, 103-112 Berkeley: Berkeley Linguistics Society.
- Brentari, Diane. (in press). *A Prosodic Model of American Sign Language Phonology*. MIT Press.
- Friedman, Lynn. (1977). Formational properties of American Sign Language. In L. Friedman (Ed.), *On the other hand: New perspectives in American Sign Language*, 13-56. New York: Academic Press.

- van der Hulst, Harry. (1995). The composition of handshapes. *University of Trondheim, Working Papers in Linguistics*, 1-18. Dragvoll, Norway.
- Itô, Junko and Armin Mester. (1995a). Japanese Phonology. In J. Goldsmith (Ed.), *A Handbook of Phonological Theory*, 817-838. Oxford: Basil Blackwell.
- Itô, Junko and Armin Mester. (1995b). The Core-Periphery Structure of the Lexicon and Constraints on Reranking. In J. Beckman, et al. (Eds.) *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory*, 181-210. University of Massachusetts, Amherst: Graduate Linguistic Student Association.
- Mandel, Mark A. (1981). *Phonotactics and morphophonology in American Sign Language*. Berkeley, CA: University of California dissertation.
- Padden, Carol. (1988). *Interaction of morphology and syntax in American Sign Language*. New York: Garland Press.
- Padden, Carol. (in press). The ASL lexicon. *International Review of Sign Linguistics*.
- Padden, Carol, and David Perlmutter. (1987). American Sign Language and the architecture of phonological theory. *Natural Language and Linguistic Theory* 5, 335-375.
- Perlmutter, David. (1992). Sonority and syllable structure in American Sign Language. *Linguistic Inquiry* 23, 407-442.
- Sandler, Wendy. (1989). *Phonological representation of the sign*. Dordrecht: Foris.
- Stevens, Kenneth, Samuel Jay Keyser, and Haruko Kawasaki. (1986). Toward a phonetic and phonological theory of redundant features. In J. Perkell, and D. Klatt. *Invariance and variability in speech processes*, 426-449. Hillsdale, NJ: Lawrence Erlbaum.
- Stevens, Kenneth, and Samuel Jay Keyser. (1989). Primary features and their enhancement in consonants. *Language* 65, 81-106.
- Supalla, Samuel. (1992). *The book of name signs*. San Diego, CA: Dawn Sign Press.
- Supalla, Ted. (1982). *Structure and acquisition of verbs of motion and location in American Sign Language*. Ph.D. dissertation, University of California, San Diego.

The Role of the Determiner in the Semantic Interpretation of NP-Type Nouns

Tracey C. Burns and Nancy N. Soja
Northeastern University

1. Introduction

The problem of how children constrain the possible meanings of new words such that they can successfully make word to object mappings has received much attention in word learning research. One area that has not traditionally received as much attention is the task a child is faced with when a word s/he knows has more than one meaning in any given context. This is a real problem for a young learner, since many words have more than one sense and the situational information may not be sufficient to distinguish between all possible meanings. Recent research on pragmatic functions such as joint attention, eye gaze (Baldwin, 1993), and notions of intentionality (Clark, in press) has all begun to study possible cues a young child could utilize in order to decipher which particular meaning is intended by a speaker. Our research has focused on a different sort of information a child could use to restrict possible word meanings to the intended one: the syntactic frame in which a word appears. The different senses of any given word are often systematically related to the syntactic contexts in which they appear: For example, think of the child whose task it is to determine the meaning of the nominal *chicken*, and look at the following two sentences:

"I had chicken."

"I had a chicken."

In the first sentence, *chicken* is a mass noun that denotes non-individuated stuff (i.e., a portion of chicken), and the speaker is most likely referring to his/her lunch order. Other than a portion, there can be no other interpretation of the noun in this sentence construction. In the second sentence, however, *chicken* is a count noun that denotes an individuated entity (namely, an entire chicken). In this sentence, although you can still interpret the nominal as a meal (albeit a very large one), there is an alternative interpretation possible: One could interpret Sentence 2 to mean that the speaker, in the past, possessed

this entity as a pet. In this sense there is a direct and meaningful change in the entity referred to in the two sentences, and this change in meaning is signaled by the sentence construction.

2. NP-type nouns

This issue is particularly significant implications for a type of noun called NP-type nouns, which are nominals that alternate between count noun and lexical noun phrase constructions with resulting changes in their semantic interpretation. Identified by Soja (1994), NP-type nouns include such nominals as *church*, *school*, and *camp*. Traditionally thought of as count nouns, they do have a count noun construction, but they differ from count nouns in several important ways. Count nouns (e.g., *building*, *table*) can be modified by count selective determiners such as *a* or *another*, and cannot be used without determiners in the singular, as we can see in the following example:

She is at a building.
* She is at building.

In contrast, NP-type nouns such as *camp* and *church* can be used both with count specific determiners and without any determiner (or bare):

She is at a camp/church.
She is at camp/church.

Likewise, you can say both "She is at a school" and "She is at school". Clearly, these nouns are not just traditional count nouns, since they can be used bare in the singular. As many of you know, the counterpart to count nouns are mass nouns. Mass nouns (e.g., *sand*) can be used in the singular without a determiner (e.g., "She likes sand"). Perhaps, then, you are thinking that NP-type nouns like *church* and *camp* are nouns that simply alternate between count and mass noun readings.

Unfortunately that is not the case, for at least two reasons:

1) mass nouns like *sand* and *mud* can be modified by mass selective determiners such as *some* or *much*. NP-type nouns do not take such modifiers, as we see in the following example:

She has sand.
She has some/much sand.

She is at camp.

*She is at some/much camp.

And 2) mass nouns can be modified prenominally while NP-type nouns cannot:

This is lettuce.

This is expensive lettuce.

She is at camp.

*She is at expensive camp.

Clearly NP-type nouns, when used bare, are not acting as mass nouns. They are also not traditional count nouns, as pointed out earlier, even though they are commonly categorized as such. Where does that leave us? We argue that NP-type nouns are in fact a third, new kind of noun that is distinct from both count and mass nouns syntactically and semantically. Syntactically, NP-type nouns behave as noun phrases when used bare, hence the appellation 'NP-type noun', where NP refers to noun phrase (see Soja, 1994, for a complete discussion). To understand what we mean by this, take, for example, a common NP such as *a building*.. Here the NP is comprised of a count noun and a determiner. Such an NP cannot be modified with a prenominal adjective, but can take a predicate adjective. So in the following example Sentence A is ungrammatical because the adjective is in the prenominal position, while Sentence B is grammatical because the adjective is a predicate.

A) *This is nice [a building].

B) [This building] is nice.

Similarly, an NP-type noun, when used bare, displays the same characteristics:

A) *This is nice church.

B) Church is nice.

Of course, we could always add the determiner and say "This is a nice church", but at that point *church* is not used bare and is in fact acting as a count noun in the sentence. Our claim is that NP-type nouns, when used bare, are acting as noun phrases.

Below is a partial list of NP-type nouns. All of the nouns listed, as well as many others not on the table, can be used in both a count noun construction and a noun phrase construction.

Table 1. Some NP-type nouns.

school	temple	jail
college	church	prison
campus	mass	court
class	shull	trial
camp	confession	
daycare	synagogue	
kindergarten		

3. Children's productive use of NP-type nouns

What do children know about these nouns? Soja (1994) studied the spontaneous speech of four children using the Child Language Data Exchange System (CHILDES) (MacWhinney & Snow, 1985) database. She found that the children made a clear distinction between count and NP-type nouns in that they often omitted determiners with NP-type nouns but not with singular count nouns. In addition, the analysis of the use of mass nouns with adjectives demonstrated that the children, although they often used mass nouns bare, clearly differentiated mass nouns from NP-type nouns. In this study, mass nouns were often used in conjunction with prenominal adjectives (e.g., *hard bread*). However, in no cases did the children use a prenominal adjective with a bare NP-type noun (e.g., *"at biggest school"*) (Soja, 1994, p. 277).

Importantly, each child did so as soon as s/he was reliably producing the count/mass distinction. The children thus recognized NP-type nouns as such at the same time as they were working out the count/mass distinction, providing evidence that children's earliest parceling of the nominal system includes NP-type nouns. The Soja (1994) study gives us distributional evidence that children understand that the NP-type noun category is distinct from that of count nouns, and reveals that children can use both constructions when using an NP-type noun.

4. Children's comprehension of NP-type nouns

Given that NP-type nouns have both a count noun and a noun phrase construction, we can return to the issue raised in the Introduction: Namely, that different sentence constructions can have specific semantic implications for the same nominal. Burns and Soja (in press) found that adults interpret an NP-type noun in its count noun construction (e.g., *She is in a church.*), as meaning that the woman is in the church for any reason at all, whether it be for a bingo game or to mop the floors. This use of the NP-type noun elicits the location sense of the word. In contrast, in a noun phrase construction (e.g., *She is in church.*), adults require the subject of the sentence to be in the church for mass or otherwise participating in a behaviour that is consistent with the habits and rituals associated with the institution of church. In this way, the NP construction elicits the institutional sense of the NP-type noun. These results can be combined with the results of the Soja (1994) analysis which demonstrated that children produce the syntactic distinction in their spontaneous speech to produce two research questions:

1) Are children using the different constructions in a way that is semantically coherent? That is, are they using the constructions to systematically differentiate two distinct meanings of the same word? It is possible that the pattern of use shown in Soja (1994) was simply reflecting a random use of the two different constructions without any attempt or understanding that the constructions might influence the meaning.

2) If children are using the constructions in a non-random manner, do they assign the same semantic interpretations as adults to the different constructions? That is, if the nominal is used bare, does it elicit the institutional meaning? And if it is used with a determiner, does it elicit the locational meaning? Even if the children were using the two constructions in a systematic manner in order to differentiate between two meanings of the same word, it does not necessarily follow that they assign the same interpretations to the two constructions as adults do.

We set out to test these questions in a set of Experiments, two of which will be reported today.

5. Experiment 1

Procedure

We tested 32 four year olds (mean age = 4;3) by showing them pairs of pictures where one member of the pair depicted the institutional sense of an NP-type noun and the other member of the pair depicted the locational sense of the same NP-type noun. There were eight such pairs of pictures in all (depicting the NP-type nouns *church*, *school*, *kindergarten*, *daycare*, *camp*, *jail*, *TV*, and *bed*). Here, for example, is the pair of pictures used to test the interpretation of the NP-type noun *church*.

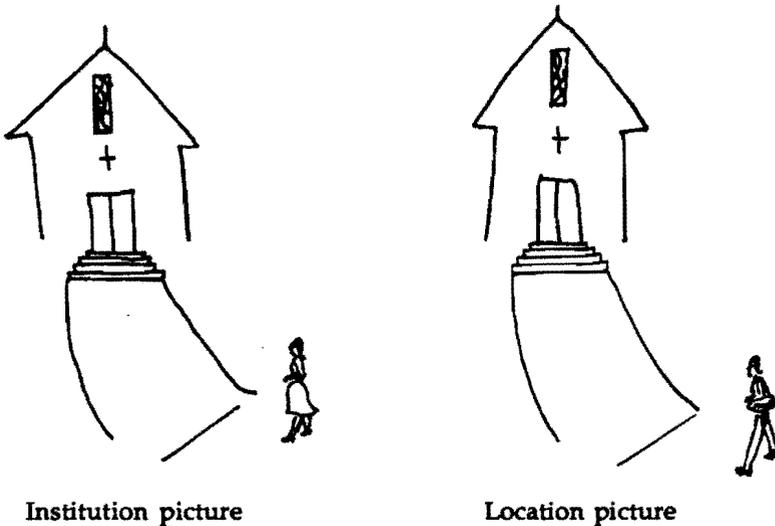


Figure 1. Sample pictures for Experiment 1

Children were shown both pictures at the same time and were read short descriptions of the scenes. For this example, they would be the following two descriptions (description A refers to the institution picture; description B the location picture):

- A) "This is a church. This is a lady. She prays here every week."
- B) "This is a church. This is a mail carrier. She delivers mail here every week."

After hearing the descriptions, children were given a test sentence (uttered by a puppet) and asked to point to the picture that best corresponded to the test sentence. The test sentences were paired so that they could either use the NP-type noun bare or with a determiner, counterbalanced across subjects. For this example, half of the children were given the test sentence "Who is going to church?", where the determiner was absent, and the other half of the children were given the test sentence with the determiner present. That is, those children would hear: "Who is going to a church?". The children were then asked to make a forced choice judgment and point to the picture that best corresponded to the test sentence they had just heard. The order of presentation of picture and type of sentence were counterbalanced across subjects.

If the children use the different constructions to express distinct meanings in a systematic fashion, we would predict that they should (minimally) treat the two sentences differently. In addition, if the children assign the same semantic interpretation as adults do to the two constructions, then it is crucial that they choose only the institutional picture when given a sentence that uses the nominal bare. That is, it is unacceptable for them to choose the location picture when they hear the nominal used bare. There is a linguistic constraint operating here: when hearing an NP-type noun used without a determiner, children cannot infer that it refers to a location. In contrast, the opposite constraint does NOT hold: It is not necessary that children choose the location picture when given the sentence containing the determiner. This is because the sentence with the determiner does refer to a location, but notice that the actual physical building of the church is present in both of the two pictures. Thus although we label one picture the institution picture and the other the location picture, they would be more properly labeled the institution/location picture and the location picture. In any case, we can make two predictions: 1) children should treat the two sentence types differently; and 2) children should choose only the institutional picture when given the sentence with the nominal used bare.

Results and Discussion

The responses to the two sentence types differed significantly from each other ($F(1,31) = 17.96, p < .0004$), demonstrating that the four year olds did not use the two constructions randomly, but instead were able to differentiate between them and use them in a semantically coherent manner. In addition, when children were presented with the sentence

in which the nominal was used bare, children were significantly more likely than chance to choose the institutional picture (see Figure 2).

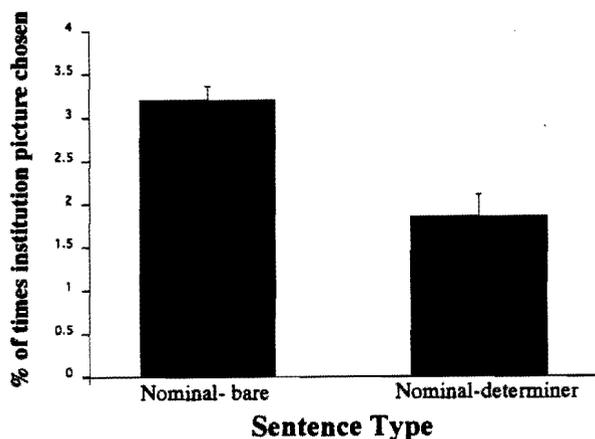


Figure 2. Results of Experiment 1

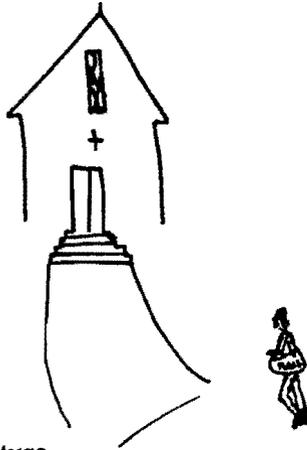
This was the crucial case, since children given the bare use of the nominal should have been constrained from choosing the location picture. In contrast, when presented with the sentence in which the nominal was used with the determiner, the children did not differ significantly from chance responding. This is not unexpected since (as already mentioned) both pictures actually portrayed the location, and therefore children could have chosen either of the two pictures and still have given the location interpretation to the sentence.

The results of Experiment 1 show that children are not random in their use of the two constructions, but instead use them in a semantically coherent way. In addition, children have the same interpretation of the two constructions as adults do: When the nominal is used bare they assign an institution interpretation, and when it is used with a determiner they assign a location interpretation.

6. Experiment 2

Procedure

In Experiment 2 we set out to replicate Experiment 1 using the same materials in a different paradigm. We used 16 four year olds (mean age = 4;3). In this paradigm, each child was shown only one of the pictures at a time and given its description. Below is the scenario that would accompany the location picture of *church*:



Location picture

Figure 3. Sample picture for Experiment 2

"This is a church. This is a mail carrier. She delivers mail here every week."

The children would then be introduced to two puppets, who would each utter one member of the sentence pair, as illustrated in the following example:

Puppet 1: "She is going to a church"

Puppet 2: "She is going to church"

The children's task was to indicate which puppet spoke correctly. They were given a series of practice trials to acquaint them with the procedure.

The institutional picture for *church* was also given to the same subject, but in a different trial. In this trial, the same two sentences would be uttered by the puppets. Order of presentation of picture, sentence, and puppet was counterbalanced across subjects. The children in Experiment 2 were given the same eight sets of pictures and sentences as those in Experiment 1.

We would predict that children should (minimally) treat the two types of pictures differently from each other. In addition, there is one crucial condition, but this time it is reversed since our paradigm is reversed. In Experiment 2, we would predict that children given the location picture must chose the sentence with the determiner, since the location picture can not be referred to with a bare nominal. In contrast, children shown the institution picture could choose either the sentence with the determiner or without it, since that picture not only portrays the institution, it also depicts the physical location of the building.

Results and Discussion

As in Experiment 1, there was a significant difference in the children's responses to the two picture types ($F(1,15) = 139.25, p < .0001$). In addition, children presented with the location picture were significantly more likely to choose the sentence with the determiner than would be predicted by chance (see Figure 4).

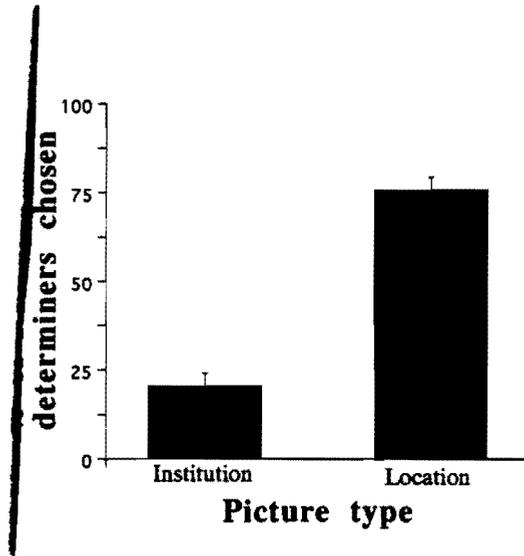


Figure 4. Results of Experiment 2

In contrast to Experiment 1, however, children in Experiment 2 shown the institution picture were also significantly more likely than chance to choose the sentence with the bare use of the nominal. An obvious question: Why do we see the split in responding here but not in Experiment 1? One possibility is that the presentation of the two sentences simultaneously made the syntactic contrast that much more salient to the four year olds, and so they were able to succeed not only at the task that had a constraint, but also be more likely to remember the preferred reading for the institutional picture. That is, given the contrast of sentences, the location aspect of the institution picture might not be obvious to the children. This is simply speculation, however, and it is not clear why the disparity of results in this condition was obtained. In any case, the difference, although interesting, does not affect the conclusions that were drawn from Experiment 1, namely, that the children treat the two sentence constructions in different yet semantically coherent fashion, and that they assign the same two semantic interpretations as adults do to the two constructions.

7. Conclusion

The two Experiments described above demonstrate that four year old children understand that the two constructions of NP-type nouns have distinct semantic interpretations associated with them. In addition, the children use the two constructions in a systematic manner that corresponds to adult usage. These experiments confirm that children understand the role of the determiner in the semantic interpretation of NP-type nouns.

The results of the present studies provide evidence that children can and do use the syntax to infer the meaning of a word in a sentence. This ability helps the child solve the problem of understanding which meaning is intended when a word s/he knows that has more than one meaning is used in an ambiguous context. Our research can be combined with recent research on pragmatic functions such as joint attention, eye gaze, and notions of intentionality in order to begin to describe possible cues a young child could utilize in order to succeed at understanding when a particular meaning of a given word is intended by a speaker. It is clear that young children can use even very subtle syntactic cues in order to help them with such tasks.

8. Acknowledgments

* Thanks to John J. Kim for helpful comments and thoughtful questions.

9. References

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20, 395-418.

Burns, Tracey C., and Soja, Nancy N. (in press). The role of the determiner in the interpretation of NP-type nouns. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, and A. Zukowski (Eds.), *Proceedings from the 21st Annual Boston University Conference on Language Development*, Somerville, MA: Cascadilla Press.

Clark, E. (in press). Speaker perspective, lexical acquisition, and semantic constraints. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, and A. Zukowski (Eds.), *Proceedings from the 21st Annual Boston University Conference on Language Development*, Somerville, MA: Cascadilla Press.

MacWhinney, B., and Snow, C. (1985). The child language data exchange system. *Journal of Child Language*, 12, 271-296.

Soja, N. N. (1994). Evidence for a distinct kind of noun. *Cognition*, 51, 267-284.

Optimizing Structure in Context: The Case of German Scrambling*

Hye-Won Choi
University of Southern California

1 Introduction

This paper examines the relationship between syntactic structure and semantic and informational meaning of language focusing on the scrambling phenomena in German. I pursue this issue from the perspective that different ordering possibilities are motivated and constrained by interactions between syntactic, semantic, and pragmatic principles of this language. In particular, I utilize Optimality Theory (Prince and Smolensky in press, Grimshaw in press) to demonstrate how these principles interact and resolve conflicts among one another to yield the “optimal” output, i.e., a sentence with a particular word order, in a given context.

2 Semantic and Discourse Effects

It has been observed that there are certain semantic and discourse-related effects associated with scrambling. These effects can largely be summarized in two generalizations: one is a focus effect (Lenerz 1977, Abraham 1986, Moltmann 1990, Webelhuth 1992) and the other is a definiteness or specificity effect (Mahajan 1990, Moltmann 1990, de Hoop 1992, Diesing 1994). Interestingly, however, both of these generalizations are challenged by a third effect, i.e., a ‘contrastive focus’ effect (Lenerz 1977, Moltmann 1990, Abraham 1986). I’ll discuss review the basic facts in this section.

*I am grateful to Joan Bresnan, Peter Sells, Henriëtte de Swart, and Thomas Wasow for their insightful comments and discussions on this and earlier versions of the paper. I also thank the audience at the 1996 Western Conference on Linguistics (WECOL) meeting for their helpful comments and questions.

2.1 Focus Effects

Scrambling exhibits a focus constraint or what I call the ‘anti-focus’ effect, which can be roughly summarized as “the scrambled element must necessarily be unfocused or topical” (Webelhuth 1992:194), and thus a focused element cannot scramble. This focus effect is well demonstrated in the contrast between examples (1) and (2) (Lernerz 1977:43). The question and answer pairs here are used to detect what the focused element is in each sentence. (The focused elements are capitalized in the examples.)

- (1) a. Wem hast du das Geld gegeben?
whom have you the money(Acc) given
‘Who did you give the money?’
- b. Ich habe dem KASSIERER das Geld gegeben.
I have the cashier(Dat) the money(Acc) given
‘I gave the cashier the money.’
- b’. Ich habe das Geld dem KASSIERER gegeben.
I have the money(Acc) the cashier(Dat) given
‘I have the money to the cashier.’
- (2) a. Was hast du dem Kassierer gegeben?
what have you the cashier(Dat) given
‘What did you give the cashier?’
- b. Ich habe dem Kassierer das GELD gegeben.
I have the cashier(Dat) the money(Acc) given
‘I gave the cashier the money.’
- b’. ?*Ich habe das GELD dem Kassierer gegeben.
I have the money(Acc) the cashier(Dat) given
‘I gave the money to the cashier.’

In (1), the dative object NP *dem Kassierer* ‘the cashier’ is the focus, and in this case, the accusative object NP *das Geld* ‘the money’, which is not focused, can easily scramble as shown in (1b’). In (2), on the other hand, the accusative object NP *das Geld* is the focus of the sentence, and the accusative object *das Geld*, which is now focused, cannot scramble, as illustrated in (2b’).

We can summarize the generalization here that a scrambled phrase should be defocused and may not be interpreted as new information. However, this generalization is undermined by another focus effect, i.e., a contrastive focus effect. As noted by Lernerz (1977), Abraham (1986), and Moltmann (1990), a contrastively focused element can actually scramble.

- (3) a. WAS hast du dem Mann gegeben? Die Zeitung?
 what have you the man(Dat) given the newspaper
 'WHAT did you give the cashier? The newspaper?'
- b. Ich habe das BUCH dem Mann gegeben (nicht die ZEITUNG).
 I have the book(Acc) the man(Dat) given not the newspaper
 'I gave the book to the man (not the newspaper).'

In example (3), the accusative object *das Buch* is the focus, but it is contrastively focused in this case. As shown in (3b), the focused object *das Buch* can scramble over the dative object. This example is in contrast with the previous example (2b').

2.2 Specificity Effect

A 'definiteness' or 'specificity' effect is another interpretation-related effect associated with scrambling. It can be summarized as 'a scrambled element should be definite or specific', and thus a nonspecific or indefinite NP cannot be in a scrambled position. This is illustrated in examples (4) and (5) (Abraham 1986:18).

- (4) a. Ich habe meinem Bruder **den Brief** geschickt.
 I have my brother(Dat) the letter(Acc) sent
 'I sent my brother the letter.'
- b. Ich habe **den Brief** meinem Bruder geschickt.
 I have the letter(Acc) my brother(Dat) sent
 'I sent the letter to my brother.'
- (5) a. Ich habe meinem Bruder **einen Brief** geschickt.
 I have my brother(Dat) a letter(Acc) sent
 'I sent my brother a letter.'
- b. *Ich habe **einen Brief** meinem Bruder geschickt.
 I have a letter(Acc) my brother(Dat) sent
 'I sent a letter to my brother.'

In (4), the definite accusative object NP, *den Brief* 'the letter', can be in a scrambled position to the left of the dative object NP *meinem Bruder* 'my brother', as shown in (4b). On the other hand, the nonspecific indefinite NP, *ein Brief* 'a letter', cannot be in that alternative position, as shown in (5b).

Interestingly, this specificity effect is also affected by contrastive focus. See example (6) (Moltmann 1990:15-16).

- (6) weil Hans ein BUCH dem Mann gegeben hat
 because Hans a book(Acc) the man(Dat) given has
 (nicht eine ZEITUNG)
 not a newspaper
 'because Hans gave a book to the man, (not a newspaper)'

The example in (6) first shows that a focused phrase, especially a contrastively focused phrase, can scramble, just as the example in (3) does. It moreover shows that, when contrastive focus is involved, the specificity effect disappears. That is, a scrambled indefinite NP retains its nonspecific reading: it can be interpreted as existential or weak (Diesing 1992, de Hoop 1992).¹ So, for example, *ein Buch* in (6) need not be interpreted as a specific book that the speaker has in mind: it can be any random book.

These new facts regarding the contrastive focus effect of scrambling lead us to conclude that neither the specificity constraint or the anti-focus constraint is the correct generalization to describe the semantic or discourse effect associated with scrambling. Furthermore, the existence of a third constraint which requires information about the focality of a scrambled element in addition to its specificity, shows that the discourse and semantic effects are not unrelated and thus should not be treated separately. The contrastive focus effect further demands that we should more closely examine discourse notions such as topic, focus, and contrastive focus, because these certainly seem to be crucial factors in determining the scramblability of a phrase. I will investigate this issue in section 3.

3 Information Structure

Following Halliday (1967), Vallduví (1992), and Lambrecht (1994), I will call the component which deals with topic and focus phenomena 'information structure'. The information structure of a sentence, roughly speaking, is a reflection of discourse-contextual information in that sentence. Depending on how it relates to discourse, a sentence can have various information structural descriptions.

Unlike the traditional binary divisions of the information field such as topic and comment, background/presupposition and focus, or theme and rheme, Vallduví (1992) proposes a ternary division of information structure. He first partitions a sentence into focus and ground, which is roughly parallel to given-new or topic-focus division, and then further partitions the ground, i.e. old or given information, into link and tail, as illustrated in (7). This subdivision

¹It seems that speakers vary in accepting the sentences in (6). See Choi (1996:Ch.3) for discussion on speaker variation.

of the ground information is crucial to his account of word order in Catalan: according to him, link elements are left-detached and tail elements are right-detached in Catalan (Vallduví 1992:Ch.5).

- (7) $S = \{\text{focus, ground}\}$
 $\text{ground} = \{\text{link, tail}\}$

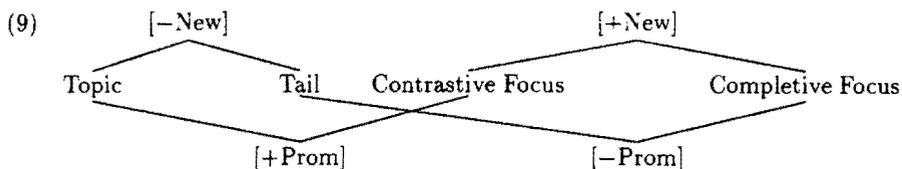
To better explain the discourse effects of scrambling in German, I propose a further subdivision also in the new or informative part of the information structure, namely, to completive focus and contrastive focus, as demonstrated in (8). Recall that the two focus-related effects are contradicting each other: the anti-focus effect says that only a nonfocused or topical element can scramble and a focused element cannot, while the contrastive focus effect says that a (contrastively) focused element can actually scramble. So, obviously, there is more than one type of focus involved here.

- (8) $\text{focus} = \{\text{completive focus, contrastive focus}\}$

Contrastive focus has often been noted as distinguishable from a purely new-information type of focus in the literature. A contrastively focused element is often compared with or even opposed to some other element, whether this comparison or opposition is explicit or implicit, or stated or predicted. Thus, a set of 'alternatives' (cf. C-Set in Rooth (1992)) is evoked. And the existence of alternatives makes the currently focused item 'prominent' so that the focused item gets 'emphasis' and often creates a 'counterexpectation' effect.

With the further division in focus added to the system, we now have four different types of information units, i.e. topic, tail, completive focus, and finally contrastive focus. As shown in (7), link (or topic) and tail are grouped together as ground materials, i.e. old or given information in discourse, while completive focus and contrastive focus are classified as focus materials, i.e. new or novel information. That is, the characteristic feature which distinguishes topic and tail on the one hand, and completive focus and contrastive focus on the other, is the 'discourse-newness'. I will call this feature [New]. Thus, topic and tail are marked [-New] and completive focus and contrastive focus are marked [+New]. Then, as I proposed in (7), both ground and focus can be further partitioned. I propose that the distinctive feature here is 'discourse-prominence', which will be called [Prom]. Topic and contrastive focus are prominent while tail and completive focus are not. So, the former are [+Prom] and the latter [-Prom].

The two information partitioning features [New] and [Prom], therefore, give the following cross-classification of the four information types as illustrated in (9).



Each information type is then represented by a pair of features in the new feature-based system, which is illustrated in (10).

(10)

	Topic	Contrastive Focus	Tail	Completive Focus
Prom	+	+	−	−
New	−	+	−	+

One of the advantages of this feature-based information structure is that it can crossrefer to more than one distinct informational type. For example, topic and tail can be grouped together as being [−New], as Vallduví does by calling them ‘ground’. Also, we can crossrefer to topic and contrastive focus together as [+Prom] elements. This crosscutting property of the system is absolutely crucial in explaining the discourse effects of scrambling as we will see shortly. For example, the [+Prom] feature is the key to the explanation of why not only a topical, but also a contrastively-focused element easily scrambles. In a discrete primitive-based system such as Vallduví’s (1992, 1993), however, it is not easy to explain why topic and (contrastive) focus should behave alike. Also, the differences between topic and contrastive focus can easily be captured by the values of the [New] feature, i.e. the former is [−New] and the latter [+New]. This difference explains the fact that topic is more scramblable than contrastive focus.

4 Optimizing Structure in Context

Having established the information structure, I argue that scrambling or alternative word order phenomenon is in fact motivated to satisfy the given discourse information or to best instantiate the given information structure. In other words, scrambling is a result of interactions between the discourse-contextual information, which forms the information structure of a sentence, and the syntactic information, which builds the basic phrase structure. In this section, I’ll derive the focus and specificity effects discussed in section 2 by means of Optimality-Theoretic interactions of syntactic, discourse, and also semantic information in the grammar. Let me first briefly review the core ideas of Optimality Theory.

4.1 Optimality Theory

In Optimality Theory (Prince and Smolensky in press, Grimshaw in press), a grammar is a function, mapping each linguistic input to its correct structural description or 'parse', or output. Inputs consist of raw materials from which the candidate outputs are built. In syntax, inputs are usually taken to be some skeletal structure containing predicate-argument information (Grimshaw in press, Bresnan 1996), tense/aspect specification, and some semantic and pragmatic information (Grimshaw in press, Grimshaw and Samek-Lodovici 1995).

I argue that each lexical element in the input is actually marked with the discourse-information features, [Prom] and [New]. (11) is an example of input, in which the subject *Hans* is topic, the dative object *dem Schüler* is completive focus, and finally the accusative object *das Buch* is tail.

(11) Input:

geben(x,y,z):	x=Hans;	y=dem Schüler;	z=das Buch;	tense=Past
	-New	+New	-New	
	+Prom	-Prom	-Prom	

Of course, the discourse feature marking in the input is not unrestricted: certain elements cannot take certain features due to their morphological, semantic, or other properties. I will argue later in section 4.3 that these feature-marking restrictions on input representations are responsible for the specificity effect of scrambling.

For every input, GEN (for "generator"), a crosslinguistically universal function, generates the universe of possible candidate outputs, i.e., the *candidate set*. I take the various scrambled structures of a sentence to be the competing candidates. For the input (11), all 6 alternative orders of arguments are possible, as shown in (12).

(12) Candidates:

a.	daß	Hans	dem Schüler	das Buch	gegeben	hat
b.	daß	Hans	das Buch	dem Schüler	gegeben	hat
c.	daß	dem Schüler	Hans	das Buch	gegeben	hat
d.	daß	das Buch	Hans	dem Schüler	gegeben	hat
e.	daß	dem Schüler	das Buch	Hans	gegeben	hat
f.	daß	das Buch	dem Schüler	Hans	gegeben	hat

Universal Grammar also provides a set of 'well-formedness' constraints on outputs. Unlike 'principles' of more familiar type in the previous theories, the constraints in OT are *violable* and *ranked*. For the scrambling data in German,

I propose two sets of constraints in (13) and (14), and the constraint ranking in (15). The first set of constraints, ‘information structuring’ constraints urge each element of a sentence to be aligned according to its information status. The second set of constraints, ‘canonical structure’ constraints require that each element be aligned according to its syntactic information such as argument structural information or grammatical functional information.

(13) Information Structuring Constraints:

- a. NEW: A [-New] element should precede a [+New] element.
- b. PROM: A [+Prom] element should precede a [-Prom] element.

The first constraint NEW in (13) immediately explains the fact that a ground element, topic or tail, which is [-New] in our information structure, can be out of its canonical position, unless it already precedes the [+New] elements in the sentence in the canonical configuration. On the other hand, the second constraint PROM explains that contrastive focus as well as topic can scramble out of its base position. These information structuring constraints NEW and PROM, working together, yield the following consequences. A completive focus, which is [+New,-Prom], is the least likely element to be out of its canonical position because neither NEW nor PROM motivates it to scramble. Also, they yield another consequence that topic, which is [-New,+Prom], is most likely to scramble because both NEW and PROM endorse its scrambling. Contrastive focus is an interesting case. Its [+New] feature discourages its scrambling, but its [+Prom] feature encourages it. Actually, PROM is ranked higher than NEW in German, as illustrated in (15). This ranking makes contrastive focus quite a scramblable entity.

As expected, alternative ordering is also restricted by the default phrase structural descriptions, yielded by the canonical structure constraints. These constraints are responsible for the so-called default or unmarked order, which is [subject-dative object-accusative object] in German.

(14) Canonical Structure Constraints:

- a. CN1: Subject should c-command non-Subjects.
- b. CN2: Non-Subject functions align reversely according to the functional hierarchy.
(Subject > Acc-Object > Dat-Object > Oblique > Adjunct)

The first constraint CN1 is a constraint about subject vs. non-subject elements, and the second constraint CN2 is one about non-subject elements. As indicated in the ranking in (15), I propose CN1 is ranked higher than CN2. This is to capture the fact that scrambling over subject is much harder than that over non-subject in German. Because of this ranking, a candidate in which subject is scrambled over by other elements is more severely penalized.

(15) Constraint Ranking (German):

$$\text{PROM} \gg \text{CN1} \gg \left\{ \begin{array}{l} \text{NEW} \\ \text{CN2} \end{array} \right\}$$

The grammatical output is the one that satisfies the constraints as best as it can, or *optimally*. Given the ranking of the constraints, the optimal output is the form which, for every pairwise competition involving it, best satisfies the highest-ranking constraint on which the competitors conflict.

4.2 Syntax/Discourse Interaction

Now, let us return to the scrambling examples in section 2 and see how the current OT account explains the data. Before we do that, first think of a situation in which a sentence is presented with no discourse-contextual information. Let me call it a 'neutral' context. In this case, no element of the sentence is newer or more prominent than any other element. Therefore, no element is marked with discourse features in the input as shown in (16). This input feeds the EVAL process in (17). (Optimal outputs in tableaux are represented with a smiling face ☺.)

(16) Neutral Context:

Hans dem Schüler das Buch
[] [] []

(17)

	CANDIDATES			PM	CN1	NW	CN2
☺a.	Hans	dem Schüler	das Buch				
b.	Hans	das Buch	dem Schüler				*
c.	dem Schüler	Hans	das Buch		*		
d.	das Buch	Hans	dem Schüler		*		*
e.	dem Schüler	das Buch	Hans		**		
f.	das Buch	dem Schüler	Hans		**		*

First look at the information structuring constraints, PROM and NEW. As expected, neither constraint applies in this situation. This indicates that unless a difference in the feature marking exists in the input, the information structuring constraints play no role with respect to word order. Hence, the canonical structure constraints CN1 and CN2 choose the default word order in (17a) as the best output. That is the only candidate which does not violate any constraint since the subject *c*-commands all the non-subject elements and the dative object *c*-commands the accusative object.

Now, let us consider how focus effects are derived in this framework. Recall that example (1) contrasts with example (2) in that the nonfocused accusative object in (1) can either stay in the base position or scramble over the dative object, whereas the focused accusative object in (2) should necessarily stay in the base position and cannot scramble. With no other linguistic clues, the dative object in (1) and the accusative object in (2) are completive focus respectively, and thus marked [+New,-Prom] in our information structure. And the rest of the elements are simply tail, marked [-New,-Prom].

First, look at tableau (19) for the input in (18).

(18) $y = \text{CompFocus}(\text{ex.}(1))$:

Hans dem Schüler das Buch
[-New,-Prom] [+New,-Prom] [-New,-Prom]

(19)

	CANDIDATES			PM	CN1	NW	CN2
⊖ a.	Hans	dem SCHÜLER	das Buch			*	
⊖ b.	Hans	das Buch	dem SCHÜLER				*
c.	dem SCHÜLER	Hans	das Buch		*	**	
d.	das Buch	Hans	dem SCHÜLER		*		*
e.	dem SCHÜLER	das Buch	Hans		**	**	
f.	das Buch	dem SCHÜLER	Hans		**	*	*

In this case, the PROM constraint would not play any role because every element has the same status with respect to [Prom]. Therefore, NEW and CN 1 and 2 together decide on the optimal output. The violation pattern for CN1 and CN2 for each output is just like the one in the tableau for the neutral context example, and it will remain the same for all the other tableaux too, because the violations of the phrase structural constraints are not context-sensitive. The candidates (c) through (f) are immediately out of consideration because they violate the constraint CN1 which is higher than NEW and CN2. With respect to the NEW constraint, a violation occurs whenever a [+New] element precedes a [-New] element. The first output (a) incurs one violation mark because the focus *dem Schüler* precedes the tail *das Buch*. The second candidate (b) does not violate the NEW constraint, but it instead violates the constraint CN2 because the accusative object *das Buch* precedes the dative object *dem Schüler* in this case. Note that NEW and CN2 are not ranked with respect to each other or both rankings are possible between the two. The result is that both (a) and (b) are optimal outputs in this context. This is

exactly what needs to be captured. Recall that in example (1), scrambling of the accusative object is optional. This optionality of scrambling is explained by the rankless competition between the discourse constraint NEW and the syntactic constraint CN2.

Now compare this example with the one illustrated in (20) and (21). Note that the optionality shown in (1) disappears in example (2). We have observed the ungrammaticality of (2b') as the 'anti-focus' effect.

(20) z=CompFocus (ex.(2)):

Hans dem Schüler das Buch
 [-New,-Prom] [-New,-Prom] [+New,-Prom]

(21)

	CANDIDATES			PM	CN1	NW	CN2
a.	Hans	dem Schüler	das BUCH				
b.	Hans	das BUCH	dem Schüler			*	*
c.	dem Schüler	Hans	das BUCH		*		
d.	das BUCH	Hans	dem Schüler		*	**	*
e.	dem Schüler	das BUCH	Hans		**	*	
f.	das BUCH	dem Schüler	Hans		**	**	*

Now the accusative object is the focus in this case, unlike that in (1). The PROM constraint does not apply in this case either, because all the elements are marked [-Prom]. Therefore, the decision is to be made by the interaction between NEW and CN 1 and 2. Again, CN1, which is higher than CN2 and NEW, quite easily narrows the choice down to (a) and (b), because all the others violate this high constraint, and thus are quickly eliminated from the competition. The NEW constraint demands that *das Buch*, which is [+New], follow other elements. Between (a) and (b), (b) is discarded because here the [+New] element, *das Buch*, precedes a [-New] element, *dem Schüler*, causing a violation of the NEW constraint. (b) also violates CN2 because the accusative object *das Buch* precedes the dative object *dem Schüler*. Therefore, (a) is the optimal output. And only this output is grammatical in this context. This result matches the fact that the question context in (2) has only one possible answer, and thus naturally captures the anti-focus effect shown in (2b'). Candidate (b) in (21), which equals (2b'), does not surface as optimal in this case unlike candidate (b) in (19).

The current analysis easily explains the contrastive focus effect as well. Look at the example (3) again. As argued in section 3, a contrastive focus is represented as [+New,+Prom] in the input, i.e. as prominent new information.

Therefore, the input will be marked as in (22).²

(22) $z=ContFocus$ (ex.(3)):

Hans dem Schüler das Buch
 [-New,+Prom] [-New,-Prom] [+New,+Prom]

As mentioned above, a contrastive focus is a very puzzling case because while its newness ([+New]) discourages it from scrambling, its prominence ([+Prom]) encourages it to do so, given the information structuring constraints proposed in this section. As expected, [+Prom] wins, and thus the focused phrase scrambles, because the PROM constraint outranks the NEW constraint in this language. This is illustrated in the tableau in (23).

(23)

	CANDIDATES			PM	CN1	NW	CN2
a.	Hans	dem Schüler	das BUCH	*			
☺b.	Hans	das BUCH	dem Schüler			*	*
c.	dem Schüler	Hans	das BUCH	**	*		
d.	das BUCH	Hans	dem Schüler		*	**	*
e.	dem Schüler	das BUCH	Hans	**	**	*	
f.	das BUCH	dem Schüler	Hans	*	**	**	*

In candidate (a), the contrastive focus *das Buch* stays in situ and thus violates PROM although it satisfies NEW by doing so. In contrast, candidate (b) satisfies PROM by scrambling *das Buch* although it violates NEW and CN2. However, since PROM outranks NEW and CN2, (b) wins over (a). Candidate (d) also satisfies the highest constraint PROM, but it violates the next highest constraint CN1. Therefore, (b) is the optimal output.

4.3 Semantics/Discourse Interaction

Finally, let us look at the specificity effect case. Recall the generalization that a nonspecific phrase cannot scramble. In the discussion of the information structure, I mentioned that the discourse feature marking can also be restricted. I propose that one of these restrictions is the nonspecific element's incompatibility with the discourse-oldness. That is, a nonspecific phrase cannot be marked [-New].

(24) Specificity Constraint:

A nonspecific phrase should not be [-New].

²Note that the subject *Hans* is also marked [+Prom]. I assume that unlike other elements of the sentence, the subject can easily be the topic of the sentence even without a high pitch or being presented in the *what about* phrase.

If a nonspecific phrase is introduced as new information, that is, as focus, which is often the case, the nonscramblability is easily explained. The NEW constraint prohibits a completive focus from scrambling. Thus, a completively-focused nonspecific phrase, just like a completively-focused specific phrase, cannot scramble. A problem arises when a nonspecific phrase is *not* introduced in the discourse as a focus or new information. One such context is when the focus of the sentence falls on some other element so that an indefinite NP is forced to be interpreted as a tail-like element. If a nonspecific NP is forced to be interpreted as a [-New] entity, which it cannot because of the specificity constraint introduced in (24), it constitutes another information unit with the predicate. By being part of the larger unit, it avoids the problem of being assigned the [-New] feature, although the larger unit is assigned [-New]. In other words, a nonspecific NP has an independent life in the discourse as a 'new' entity, but once it has lost its 'newness', it loses its independence, thus gets a 'predicate-modifier' type of reading (cf. de Hoop 1992).

Therefore, in this context, a nonspecific NP, the accusative object *ein Buch* in (5a), for example, is not marked with any discourse features as illustrated in (25).

- (25) $y = \text{CompFocus}$ (ex.(5), cf.(1),(4)):
 Hans dem Schüler ein Buch
 [-New,+Prom] [+New,-Prom] []

Note that *ein Buch* is not assigned a [Prom] feature either. It follows from the fact that the NP is now part of a larger information unit, and thus cannot be assigned any feature individually. This predicts that a nonspecific indefinite NP cannot be interpreted as a topic either. This seems to be a desirable result. For example, a nonspecific NP cannot be put in a *what about* phrase unless it is interpreted as generic or specific, as shown in (26).

- (26) # What about a book?

This lack of feature marking of the nonspecific object explains its nonscramblability.

- (27)

	CANDIDATES			PM	CN1	NW	CN2
∩a.	Hans	dem Schüler	ein BUCH				
b.	Hans	ein BUCH	dem Schüler				*
c.	dem Schüler	Hans	ein BUCH	*	*	*	
d.	ein BUCH	Hans	dem Schüler	*	*		*
e.	dem Schüler	ein BUCH	Hans	**	**	*	
f.	ein BUCH	dem Schüler	Hans	**	**	*	*

Since *ein Buch* is not marked [-New], it does not violate the NEW constraint even if it follows the [+New] element, *dem Schüler*, as in (27a). In addition, the phrase structural constraints CN1 and CN2 forbid it to be out of its canonical position. Consequently, the first output, the canonically ordered one, is the optimal output. And this is why the scrambled variant in (5b) is bad.

Finally, consider what happens if a nonspecific phrase is a contrastive focus. Interestingly, in this case, a nonspecific indefinite phrase such as *ein Buch* can scramble. Recall that this has been the most troublesome case for previous analyses which are either based on specificity (de Hoop 1992, Diesing 1992) or on focality (Lernerz 1977, Webelhuth 1992). This case runs against both of these accounts, i.e., 'a scrambled phrase should be specific', or 'a scrambled phrase should be unfocused', because the indefinite NP *ein Buch* in (6) is neither specific nor unfocused.

In the current account, this is not a special case. Just like the definite phrase, the indefinite NP *ein Buch* is marked [+New,+Prom] since it is contrastively focused in this context.

(28) z=ContFocus (ex.(6)):

Hans dem Schüler ein Buch
[-New,+Prom] [-New,-Prom] [+New,+Prom]

The nonspecific phrase *ein Buch*, being [+New] now, is not constrained by the SPECIFICITY constraint. Remember that the SPECIFICITY constraint only controls the cases where a nonspecific phrase is marked [-New].

(29)

	CANDIDATES			PM	CN1	NW	CN2
a.	Hans	dem Schüler	ein BUCH	*			
☺b.	Hans	ein BUCH	dem Schüler			*	*
c.	dem Schüler	Hans	ein BUCH	**	*		
d.	ein BUCH	Hans	dem Schüler		*	**	*
e.	dem Schüler	ein BUCH	Hans	**	**	*	
f.	ein BUCH	dem Schüler	Hans	*	**	**	*

The PROM constraint narrows down the competition to (b) and (d). Because (d) violates CN1, (b) is the optimal output. And this explains why a contrastively focused indefinite can retain its weak or existential reading even when it is scrambled. The nonspecific NP *ein Buch* maintains its [+New] marking (because it is not restricted by the SPECIFICITY constraint), and as a [+Prom] element, it can be scrambled.

5 Conclusion

The current OT analysis nicely captures the semantic and discourse effects including the problematic contrastive focus case, using the feature-based information structure as the major part of input representation. To summarize, I have argued that each scrambled variant is the best structural description of a particular discourse-contextual information in the input, and this is constrained by a small number of syntactic and discourse constraints, which is further restrained by a semantic constraint. OT provides a useful tool to deal with word order phenomena in German, which I argue are a consequence of interactions of the constraints from different modules of grammar, because it is designed to deal with potentially highly conflicting constraints.

Bibliography

- Abraham, Werner. 1986. Word order in the middle field of the German sentence. In Werner Abraham and Sjaak de Meij (eds.), *Topic, Focus, and Configurationality*, 15–38.
- Bresnan, Joan. 1995. Lexical-functional syntax. MS., Stanford University.
- Bresnan, Joan. 1996. Optimal syntax: Notes on projection, heads, and optimality. MS. Stanford University.
- Choi, Hye-Won. 1996. *Optimizing Structure in Context: Scrambling and Information Structure*. Doctoral dissertation, Stanford University.
- de Hoop, Helen. 1992. *Case Configuration and Noun Phrase Interpretation*. Doctoral dissertation, University of Groningen.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA, MIT Press.
- Diesing, Molly. 1994. Semantic variables and object shift. MS. Presented at the 9th Comparative Germanic Syntax Workshop, Harvard.
- Grimshaw, Jane. in press. Projection, heads, and optimality. To be published in *Linguistic Inquiry*.
- Grimshaw, Jane, and Vieri Samek-Lodovici. 1995. Optimal subjects and subject universals. Presented at the MIT Workshop on Optimality in Syntax.
- Halliday, Michael A. K. 1967. Notes on transitivity and theme in English. *Journal of Linguistics* 3, 199–244.

- Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge, Cambridge University Press.
- Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen, TBL Verlag Gunter Narr.
- Mahajan, Anoop K. 1990. *The A/A-bar Distinction and Movement Theory*. Doctoral dissertation, MIT.
- Moltmann, Friederike. 1990. Scrambling in German and the specificity effect. MS., MIT.
- Prince, Alan, and Paul Smolensky. in press. Optimality theory: Constraint interaction in generative grammar. To be published in MIT Press.
- Rooth, Mats E. 1992. A theory of focus interpretation. *Natural Language Semantics* 1, 75-116.
- Vallduví Enric. 1992. *The Informational Component*. New York, Garland.
- Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation*. New York, NY, Oxford University Press.

Explaining Coronal Underspecification
Michael Dobrovolsky
University of Calgary

0 Introduction: the Form and Substance interface

This paper is intended as a contribution to what, many years ago, was called the Form versus Substance debate—what we might today call the Phonetics/Phonology interface, or the war between phonetics and phonology.

For some scholars, of course, there is no debate. The phonologist John Goldsmith (1990: 10) has put it this way:

We . . . must remember that while phonetic reality may motivate a phonological representation, it neither justifies it or ultimately explains it. Phonetic reality provides the stuff of which phonological theory provides the explanation.

The phoneticians John and Manjari Ohala have put it this way (1991: 273):

In bypassing it [feature geometry] we do not abandon the search for the causes of speech sound behavior; rather we find far more satisfying accounts in current models of speech production and perception—many of them quite formal and all of them based on first principles, not on ad hoc theoretical constructs created only for the problem at hand.

As is so often—but not always—the case in nature and in science, the final outcome will probably lie in the complex tangle of systems and subsystems, physical and cognitive, adaptive and exaptive, that fill the gap between these two extreme positions.

I take as exemplifying an "interactive" position the formulation of Rossi (in Docherty and Ladd 1992: 227):

. . . the form [phonology] imposes an organization on the parameters of the matter [phonetics] according to the constraints and the specific modes of the matter.

I would go further than Rossi and say that there are cases, such as the one I discuss in this paper, in which the "constraints and specific modes of the matter" can in fact be shown to "impose an organization" on the parameters of the form, one in which phonetic reality will serve to correct non-explanatory phonological representations, justify new ones, and offer an explanation for them.

0.1 Coronal underspecification as an illusion

Much evidence supports a phonological theory of coronal underspecification (Paradis & Prunet 1991) such as the facts that (certain) coronals are variously exempt from certain consonant harmony processes, tend to show up frequently as default consonants, and are subject to assimilation where certain other consonants are not.

To what can we attribute this patterning? In other words, how can we explain it? Stemberger and Stoel-Gammon 1991 consider it to be a form of frequency bias, and Mohanen 1991 claims that the special properties of coronals fall out from the fact that they are unmarked, or natural in some not completely defined way.

Generally, the explanation offered for coronal underspecification has been a circular one. For example, a pattern of assimilation is analyzed as implying that coronals are best represented as underspecified; they are omitted from the hierarchical feature tree, and then the representation is said to explain the patterning. I am not being facile here. Steriade 1995: 126, discussing Kiparsky's analysis of the facts of Catalan nasal assimilation (where only alveolars undergo place assimilation, while only nonalveolars trigger it), puts it this way: "The explanation proposed by Kiparsky is that a marking condition . . . prohibits the lexical appearance of [+coronal] . . . The consequence . . . is that coronals are placeless in the underlying representations, as well as throughout the lexical component." Avery and Rice 1989: 190, also for Catalan: "The unique patterning of the coronals in Catalan is best explained by the absence of the Coronal node . . ." and (p. 193) Chumash: "The harmony rule acts as if Coronal were absent for the stops as they do not block harmony."

I propose that coronal underspecification is essentially an illusion. Its apparent existence arises from the articulatory persistence of complex articulations [NB — not complex segments, in the phonological sense], both coronal and non-coronal, and the inherent rapidity and flexibility of the coronal articulator. This leads to a view of articulations as being more or less articulatorily complex, with coronals among the least complex and therefore the most likely to undergo the types of processes associated with coronal underspecification. I show this in the following way. In section one, I review data on tongue musculature. Section two deals with recent work on the nature of articulatory complexity and its relation to phonetic processes, especially assimilation. In section three, I propose an articulatory explanation for coronal underspecification. Section four considers the implications for a feature hierarchy of the conclusions of section one and two. To limit the scope of the paper, I will have little to say about labials.

1.0 Musculature

The tongue tip and tongue body are differentiated by virtue of the greater mass of the dorsum and also the greater articulatory control we have over the tip. The reasons for this difference are well known. I review the two tongue areas in the following sections. Note that all of the phonetic material reviewed in this paper is non-tendentious with respect to the issue of coronal underspecification: it was all created for different ends, and so we can assume a certain objectivity with respect to the question at hand.

1.1 Inherent capabilities

The tongue tip is more flexible and can operate more rapidly than both the tongue body and lips. Hudgins and Stetson 1937: 92, in Hardcastle 1976: 89, provide figures on articulator rate capabilities (non-speech). I have averaged these for presentation in Figure 1.

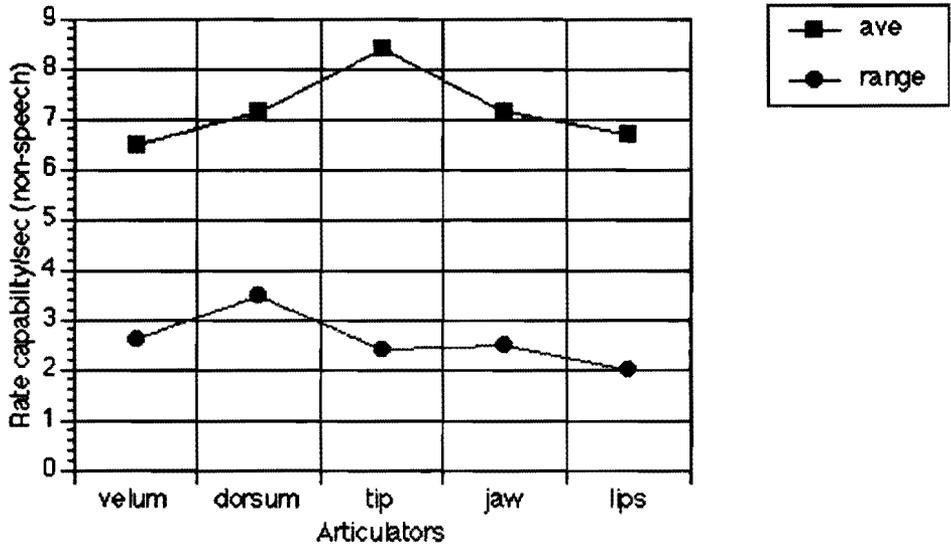


Figure 1. Articulator rate capabilities. After Hudgins and Stetson 1937, cited in Hardcastle 1976

This figure shows that the tongue tip has both the highest movement rate capability and at the same time the lowest movement range among the articulators examined. With less range to cover and a high movement rate, it's clear that the tongue can be deployed and redeployed quickly. The dorsum is a markedly slower operator and must cross a wide range of movement, at least in its maximum capability. We shall see that this rate and range capability data is reflected in real speech movement data.

1.2 Muscle movers

Inherent rate differences devolve from the activation of the movements by inherent musculature. Tongue tip movements are controlled by intrinsic tongue muscles, while dorsal activity is controlled by extrinsic musculature, as illustrated in Table 1.

muscle	type	function
<i>intrinsic</i>		
long. superior (only unpaired)	longitudinal; sheet (anterior); spread (posterior)	retracts and curls tip (pulls up); shortens, widens the tongue
long. inferior	longitudinal	pulls down and retracts tip; shortens tongue front-to-back
transversus	layers fan out towards tongue margins	draws edges of tongue upward; narrows (& protrudes tongue); grooves it
verticalis	short fibres, run downward	flattens tongue sideways; may help groove
<i>extrinsic</i>		
styloglossus	flat, narrow fan shaped sheet	retracts and elevates; may help in sulcalization
palatoglossus	downward	raises posterior section of tongue when velum fixed; also acts as palate depressor
genioglossus	flat, triangular, fan	anterior fibres: retract and depress tip (with l. inferior); posterior fibres: draw tongue forward in fixed mandible state; also elevates hyoid and larynx
hyoglossus	quadrilateral sheet; vertical fibres	draws tongue down and back; also with geniogloss and l/ inferior, pull down tip; depressor to elevator of other muscles; can act in reverse to elevate hyoid

Table 1. Intrinsic and extrinsic muscles of the tongue (Carola, Harley & Noback 1993, Hardcastle 1976, Kaplan 1960)

The major functions of the extrinsic musculature (styloglossus, palatoglossus, genioglossus) are to raise and retract the tongue body and to draw the tongue body down (hyoglossus).

The intrinsic musculature has essentially two purposes—curling and uncurling the tongue tip and raising and lowering the sides of the tongue. These movements are subject to considerable refinement as the variety of coronal and laminal articulations attests, in contrast to the rather limited variety of velar sound types. For example, eight of seventeen places of articulation proposed in Ladefoged and Maddieson 1996: 40-41 are blatantly coronal and two others (linguo-labial and interdental), involve the tongue tip, while two are dorsal). As expected, the variety of coronal-laminal articulations has its reflex in the fine-grained anatomical nature of the intrinsic musculature. As reported in Perrier et al 1996 the tongue contains numerous types of mechanoreceptors. Sensory endings progressively decrease in density from front to back (reference to Grossman 1964); muscle spindles

(reference to Cooper 1953; see also Bowman 1971), which offer feedback on stretching and contraction of the muscle as well as rate of change, are also distributed non-uniformly, in particular "in the superior longitudinal muscle near the midline and in the front third of the tongue, and in the transverse muscle in the mid region towards the lateral borders" (Perrier et al: 56). Recently, Nguyen et al 1996 have provided experimental data that supports a model in which the tongue tip and blade form an articulator that is independently controllable with respect to the dorsum.

In sum, two distinct systems are operating to provide tongue movement for speech purposes. The anterior system is highly flexible and capable of rapid movement within a relatively narrow range of articulation. In addition, the anterior and to some extent the mid portions of the tongue are highly sensitive to mechanical and tactile feedback. The dorsal system is slower and must cover a larger range of articulation. The next section shows these systems in action.

1.3 Linguistic data

Perkell 1967, a cineradiographic study of one speaker of English, provides a great deal of detailed data on a small range of utterances. Table 2 provides some times that I have measured off from the published illustrations. Measurement was done by expanding the figures on xerox copies and creating a time-scale template based on the 100 ms interval markers provided in the original figures.

<i>all times est. ms</i>	Total closure	Time to closure	Drop-off to release	Whole gesture (begin. of mvt to drop-off)
hə'pe	122.85	105.7	5.7	222.85
hə'te	114.5	42	36.8	121
hə'ke	106.4	74.4	46	133.33

Table 2. Parameters of articulatory activity for three stops. Extracted from Perkell 1967.

Observations.

1.3.1. Total closure: Labial clearly the longest; velar shortest perhaps in anticipation of following front vowel.

1.3.2. Time to closure: Labial activity very slow to begin, then rapidly accelerates; this slow initiation extends its time-to-closure. Tongue tip gesture is very abrupt, 'vertical'. Velar gesture is a sluggish curve.

1.3.3. Drop-off to release: Rapid labial dropoff is probably aided by jaw opening. For tongue tip, there is a steep rapid curve. For velum, a slow curve that starts early in the articulation.

1.3.4. Whole gesture (beginning of movement to drop-off): Nature of labial closure pads total; coronal obviously the briefest gesture; velar perhaps shorter than anticipated, but this reflects early re-deployment, which it itself may reflect "cognitive compensation."

1.3.5. More on onset and release of consonant. Perkell (p. 20) notes that the axis along which tongue height is measured in these cineradiographic images (his axis B) is further forward than the expected place of articulation for the /k/ and thus may somewhat disrupt the picture of the articulation: "because axis B probably does not pass exactly through the place of articulation for the /k/, there are no observable sharp discontinuities corresponding to consonant onset and release."

We can thus get some idea of the nature of the gestures from the data provided. Perkell notes the early onset of the dorsal gesture during the articulation of /h/ and observes that "... the motion is faster than is generally observed for vowel articulations, but because of differences between musculature and the larger mass which must be moved, this motion is somewhat slower and must begin earlier than the comparative tongue-tip gesture for postdental consonants. The approximate durations for these gestures involving tongue mass and tongue tip are 75 ms and 25 msec, respectively." (I estimated the time-to-close as 74.4 msec and 42 msec, respectively, but I am obviously not as familiar with the images as Perkell and drew my start-line earlier; his estimate is more strongly supportive of the point I am trying to make here than is my own.)

Given the inherent inertia of labial and especially velar movements, and the rapidity of tongue tip movements, we might be led to expect that a speech production system that was cognitively, if subconsciously, "aware" would anticipate and nullify or at least average out their effects. To some extent, the total closure times support this view. But the system itself cannot be entirely overcome. The fact that we can identify a phenomenon that we label coronal underspecification suggest that something about the properties of the system itself shapes phonological patterning. Otherwise, any articulation might be expected to be randomly underspecified. (To be sure, there is some evidence that velars, too, can be phonologically underspecified (Rice 1996); it has also been suggested that phonological coronal underspecification may be set from language to language (Yip 1991: 74)). This "something that shapes phonological patterning" is, in part at least, articulatory complexity.

3 A theory of articulatory complexity

Recent work by Recasens et al has suggested that the nature of articulatory complexity can be quantified. Their work attempts to define the palatal area of articulation and the nature—simple or complex—of palatal and prepalatal articulations. In the course of this work, they propose that articulatory complex segments are definable by: i. articulatory control; ii. coupling, and iii. articulator flexibility. These claims are supported by data from alternations, assimilation, and coarticulation. Let us review each one in turn; I have somewhat altered their presentation and will supplement it.

i. articulatory control. Recasens et al suggest that there is an "inverse relationship between the degree of involvement exhibited by a given articulator in the production of a particular gesture and the degree of sensitivity to coarticulatory effects." For example, [s] is more resistant to following vowel effects than, say, [n] ("little coarticulatory sensitivity to vowel-dependent effects at the tongue front").

ii. coupling. The more coupling between or among articulators, the less likely they are to be involved in assimilatory or coarticulatory activity; I call this the "take it if it's free" principle. "V-to-V effects on the tongue dorsum are larger for [n] than for [ŋ]" and, more subtly, "English alveolar [t] appears to be less resistant . . . to closure backing and tongue blade raising effects associated with adjacent [i], than is

Italian dental [t]” this explained as being due to greater tongue front / tongue dorsum coupling for the English articulation.

iii. articulator flexibility. Less flexible articulators or aspects of articulation are less likely to coarticulate. For example, laminal articulations are more resistant to coarticulation than apical articulations (ref. to Bladon and Nolan 1977).

So we can see that even within coronal articulations, not to mention across coronal-non-coronal classes, rather fine articulatory detail can be seen to be a determining factor in phonetic patterning.

To these parameters I would add

iv. inertia. Intrinsic muscular operations and muscle mass are another aspect of “articulatory complexity.”

We may also hypothesize that the inverse of the complexity parameters will produce inverse results. Thus, articulatory simplicity is defined as requiring or characterized by

- i. less control
- ii. less coupling
- iii. less flexibility
- iv. less inertia

in articulations, and will favor or give way to greater assimilation and coarticulation.

Duration as a result of greater POA contact is also said by Recasens et al to playing a role in resistance to assimilation. For the purposes of this paper, I will assume duration under (i), articulatory control.

4 Coronal underspecification: the explanation

We are now led to the heart of the explanation I propose for coronal underspecification: in the case of coronal versus non-coronal articulations, non-coronal oral articulations require—in combinations not yet well understood but clearly describable—greater area, greater coupling, more control requirements, and show greater inertia; in a word, they are articulatorily more complex and tend to be inert as a result of this. At the same time, the fact that coronal articulations—and the more tip-coronal, the better—show less contact area, less coupling, fewer control requirements, and less inertia guarantees that they are rapidly deployable and easy to release, redeploy and, above all, to replace. In other words, “pure” coronals are articulatorily simple along a number of dimensions and as such, give way to the various control, coupling, flexibility, and inertial requirements of more complex articulations.

I emphasize straightaway that I am not proposing yet another monotheoretical, linear, and unidimensional “explanation” for everything that goes on in either phonetics or phonology. If anything, I am complicating the picture with reference to another phonetic subsystem that has an effect on the outcome of anything phonological. Everything interacts, and the demands of articulatory complexity may at times be masked by conflicting subsystems. On the phonetic plane, acoustic factors obviously play a hugely significant role, as do language-specific system inventories. Finally, “yielding,” assimilation, and coarticulation are scalar and not absolute. The following section further considers this bottom-up theory of articulatory complexity and its implications for coronal underspecification.

4.1 Some illustrations

4.1.1. Catalan. A brief look at the Catalan nasal assimilation mentioned above will serve as a first illustration. Recall that in Catalan nasal assimilation, only alveolars undergo place assimilation, while only nonalveolars trigger it. From the perspective that I am taking here, this is expected *inherently*, as the nonalveolars (here, labials, palatals, velars) are more articulatorily complex than the alveolars in at least the dimension of control. The alveolar nasals need not be thought of as underspecified, but rather yield on quite natural phonetic grounds to the other articulations. What needs to be specified is the direction of assimilation—regressive for Catalan (although in the case of C₁C₂ assimilation, the regressiveness itself may be explained on acoustic grounds—see Ohala 1990). In underspecification theory, the direction of assimilation is encoded by the very position of the blank—another instance of an illusory prediction.

4.1.2. Chumash. System inventory and directionality play a role in assimilation to articulatorily complex or simple segments, as Chumash sibilant harmony shows. Beeler 1970:16 notes that “the agreement observed to exist between and among sibilants in Chumash words takes the form of a regressive assimilation, and depends upon the class of the word final sibilant; if that final is a blade consonant, then all the preceding sibilants will, in principle, be blade consonants, but if it is an apical consonant, then they will appear as apical sibilants.” (At least one prefix is an occasional exception to the generalization; this appears to be lexically determined.) An example is shown in (1), where the alternating segment is in bold face.

1. a. ksaqutinán`us ~ b. faqutinán`if
I tell him a story *story*

At the same time, alternations exist like the one in (2).

2. a. pufpel ~ b. puspelus
You hold something *You hold it for him*

Here, the harmony works to render a blade consonant apical; no trigger accounts for the *f* of 2a., which must be underlying, yet it alternates with *s* in 2b., where the final sibilant is apical. The harmony must be viewed as shaped by directionality of application that favors the simplicity/complexity status of the final sibilant. As described by Avery and Rice 1989, Chumash sibilants undergo assimilatory harmony for the feature [posterior] not through feature-changing directional assimilation, but through fusion of dependents of the Place node that show identical primary content (194). This fusion must be stipulated as right-headed on a language-specific basis. They propose that the stops are underspecified for Coronal and thus are ignored during the feature fusion operation. But is anything gained by recourse to underspecification that could not be had in other ways? Clearly, speakers of the language were sensitive to sibilants as a class; non-sibilant *t* does not emerge as its sibilant counterpart *tʃ* in these alternations. Sensitivity to the class of sibilants as targets automatically eliminates stops as candidates for the alternation. Reinforcing this auditory/acoustic underpinning is the fact that fricatives are on the whole are articulatorily more complex than stops, involving as they do finer control of the articulatory mechanism in their production. Regarding the Chumash assimilation pattern itself, I suggest that the relatively simpler control demands of a final apical articulation “relax” the requirements for the blade

articulation and, conversely, the relatively more complex requirements of the final blade articulation— greater tongue body control and tongue sulcation, for example— "demand" anticipatory preparation in a preceding apical articulation. Again, see Ohala 1990 for acoustic underpinnings in such cases.

4.1.3. *Turkic*. Turkic rounding harmony illustrates an acoustic-articulatory effect on phonological systems. Dobrovolsky 1995 notes that presence of non-high rounded vowels in non-initial syllables has posed a phonetic problem in Turkic due to their lack of acoustic fitness. I would now add to this observation that rounded vowels in general require an extra dimension of articulatory control and so are also more articulatorily complex. Speakers of Turkic languages have resolved the problem in various ways: by disallowing non-high rounded vowels in non-initial syllables in Turkish, by centralizing and reducing them in Chuvash, and by following the path of inertia in Saxa (Yakut) in permitting labial harmony throughout a word.

4.1.4. *English*. Stemberger 1991: 76 notes that in certain speech errors, "the proportion of all errors that are exchange errors (where the two consonants are reversed, as in *Homsky and Challe* for 'Chomsky and Halle') is high when one of the consonants is underspecified with respect to the other." He illustrates this point with the contrasts in the following example.

2.	<i>underspecified</i>	<i>vs</i>	<i>specified</i>
place	alveolars		labials, velars, palatoalveolars
manner	stops		nasals, fricatives
voicing	voiceless obstruents		voiced obstruents

The case seems to me that in each contrast type, the underspecified class is the articulatorily more simple one: among the stops, alveolars on the basis of control (labials, velars, palatoalveolars), coupling (velars and palatoalveolars), flexibility (velars and palatoalveolars), and inertia (velars at least); among the manners, stops on the basis of control (nasals— velar lowering in addition to the oral articulation and fricatives on the basis of more complex closures) and coupling (nasals), and voiceless obstruents on the basis of control and coupling (the laryngeal activity and its influence on the nature of articulations in the supralaryngeal vocal tract). Again, a theory of phonology that builds in the likelihood that more complex articulations will be more inert than simple ones will be able to explain these facts without recourse to underspecification.

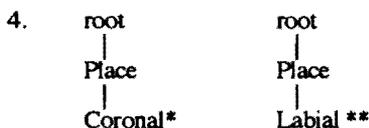
4.1.5. *Pharyngeals*. The notion of articulatory complexity is not necessarily limited to oral articulations. Outside the oral vocal tract, pharyngeals demonstrate articulatory complexity based on control, inertia and coupling. Elgendy 1995 notes that pharyngeal consonants "... prevent the anticipation of the ensuing segment to take place before the execution is completed." Alternatively, keeping the articulation in place results in pharyngealized stretches of speech. Recent laryngoscopic analysis of pharyngeals by John Esling and his colleagues at the University of Victoria (pers com) shows clearly the articulatorily complex nature of these sounds, which involve various constrictions along the length of the entire supraglottal vocal tract tube, "... beginning with glottal closure, progressing through ventricular fold closure to full aryepiglottic fold sphincter closure." Raised and lowered larynx positions add another dimension of complexity to this class of sounds.

The next question arises naturally: what are the implications of a theory of articulatory complexity for phonological representations?

underspecification that has cognitive underpinnings, since it focuses on speakers' internal representations of speech sounds.

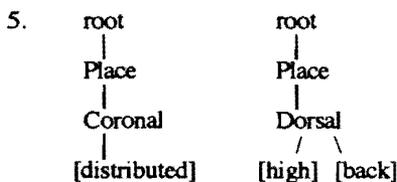
I conclude that if we are going to allow for underspecification, we are obliged to support to some extent its (cognitive) reality, and not excuse it as a metaphor for something else. But if we approach it from the perspective of articulatory complexity, how can we reflect this in the representations?

As a start, I propose that representations that incorporate a formal theory of articulatory complexity shouldn't use non-explanatory blanks. For example, in (4), I have crudely indicated the more complex segment by using an increasing number of asterisks



Where more asterisks indicate greater articulatory complexity; lesser articulatory complexity yields to greater where assimilation (a) exists and (b) is specified for directionality (as necessary).

But it's possible to incorporate articulatory complexity even more directly into the representation by emphasizing a phonetically realistic view of the feature hierarchy and providing quantitatively more feature material for more complex articulations, as long as the visual metaphor is not forced and the additional feature material is justifiable. In some cases, this phonetic realism is already present in existing hierarchies, as suggested in (5), illustrating coronal-to-velar assimilation.



Where more dependents indicate greater articulatory complexity; lesser articulatory complexity yields to greater where assimilation (a) exists and (b) is specified for directionality (as necessary).

At the present time, various feature hierarchies do not appear to require this option formally or on phonetically principled grounds, though the demands of phonological patterning have led to occasional good isomorphy between the phonetic and phonological data. For example, dorsals in at least some proposed hierarchies may contain more dependents than coronals. I believe that openly taking such a good isomorphy approach may allow a fruitful interaction between the patterning requirements of phonology and our increasing knowledge of the articulatory capabilities of the speech system. Note also in this respect Rice and Avery 1995 view first language acquisition as the emergence of (phonologically defined) complexity that is characterized by successive ramification of nodes on the feature hierarchy tree, i.e., through the elaboration of structure.

6 Conclusion for now

What we call coronal underspecification need not be thought of or represented as underspecification at all. Indeed, Boyce, Krakow, and Bell-Berti 1992: 220 claim that "even segments which lack specification by contrast criteria or by aerodynamic/articulatory criteria nevertheless exhibit characteristic articulatory positions associated with these features" and (232) "... that these positions may be obscured because of temporal constraints." Their conclusions (233) do not support a strong view of underspecification, but they do leave the door open for a continued split between phonetic and phonological approaches. I have claimed here that phonetic theory provides enough equipment to drive the engines of assimilation and default segments, and that an emerging theory of articulatory simplicity inherently does the work of a metaphor like underspecification. Some speculative options for representation have been provided. I have also argued that phonological representations that employ a hierarchical feature system should be as isomorphic with the known facts of phonetics as possible. It may not be too optimistic to argue that progressive refinement of the phonological feature system is providing just such a matchup, though it is still too early to get complacent.

Bibliography

- Avery, Peter, & Keren Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6, 179-200.
- Beeler, M. S. 1970. Sibilant harmony in Chumash. *IJAL* 36, 14-17.
- Bladon, R. A. W., and F. Nolan. 1977. A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics* 5, 185-193.
- Boyce, Suzanne E, Rena A. Krakow, and Frederika Bell-Berti. 1991. Phonological underspecification and speech motor organization. *Phonology* 8, 219-236.
- Bowman, James P., Jr. 1971. *The Muscle Spindle and Neural Control of the Tongue: Implications for Speech*. Springfield, Ill.: Charles C. Thomas.
- Carola, R., J. P. Harley, & C. R. Noback. 1993. *Human Anatomy and Physiology*. New York: McGraw Hill, Inc.
- Dobrovolsky, Michael. 1995. The phonetics of reduced vowels in Chuvash: implications for the phonology of Turkic. In Elenius, Kjell and Peter Branderud (eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 1: 62-65. Stockholm: KTH and Stockholm University.
- Elgendy, A. 1995. Testing a dynamic model of pharyngeal articulation. In Elenius, Kjell and Peter Branderud (eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 1: 354-357. Stockholm: KTH and Stockholm University.
- Goldsmith, John A. 1990. *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.
- Hardcastle, W. J. 1976. *Physiology of Speech Production*. New York: Academic Press.
- Hudgins, C. V. and R. H. Stetson. 1937. Relative speed of articulatory movements. *Archives néerlandaises de phonétique expérimentale* 13, 85-94.
- Jun, Jongho. 1996. Place assimilation as the result of conflicting perceptual and articulatory constraints. In Camacho, José, Lina Choueiri, & Mako Wanatabe (eds.), *The Proceedings of the 14th West Coast Conference on Formal Linguistics*, 221-237. Palo Alto: Center for the Study of Language and Information.

- Kaplan, Harold M. 1960. *Anatomy and Physiology of Speech*. New York: McGraw-Hill Book Company.
- Ladefoged, P., and I. Maddieson. 1996. *The Sounds of the World's Languages*. Chicago: University of Chicago Press.
- Lahiri, Aditi, and William Marslen-Wilson. 1992. Lexical processing and phonological representation. In Docherty, Gerard J., and D. Robert Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, 229-254. Cambridge: Cambridge University Press.
- Mohanan, K. P. 1991. On the bases of underspecification. *NLLT* 9, 285-325.
- Nguyen, Noël, Alain Marchal, and Alain Content. 1996. Modelling tongue-palate contact patterns in the production of speech. *Journal of Phonetics* 24, 77-97.
- Ohala, John. 1990. The phonetics and phonology of aspects of assimilation. In Kingston, John, and Mary E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 258-275. Cambridge: Cambridge University Press.
- Ohala, John, and Manjari Ohala. 1991. Reply to commentators. *Phonetica* 28 (2-4), 271-274.
- Paradis, C. and J-F. Prunet (eds.). 1991. *The Special Status of Coronals: Internal and External Evidence*. [Phonetics and Phonology, ed. S. R. Anderson and P. A. Keating, Vol. 2]. New York: Academic Press.
- Perkell, Joseph S. 1969. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study* [Research Monograph No. 53]. Cambridge, Mass.: MIT Press.
- Perrier, Pascal, Hélène Lævenbruck, and Yohan Payan. 1996. Control of tongue movements in speech: the Equilibrium Point Hypothesis perspective. *Journal of Phonetics* 24, 53-75.
- Recasens, Daniel, Edda Farnetani, Jordi Fontdevil, and Maria Dolors Pallerès. 1993. An electropalatographic study of alveolar and palatal consonants in Catalan and Italian. *Language and Speech* 36 (Pt. 2-3), 213-234.
- Rice, Keren. 1996. Default variability: the coronal-velar relationship. *NLLT* 14, 493-543.
- Rice, Keren, and Peter Avery. 1995. Variability in a deterministic model of language acquisition: a theory of segmental elaboration. In Archibald, John (ed.), *Phonological Acquisition and Phonological Theory*, 23-42. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Stemberger, Joseph Paul. 1991. Radical underspecification in language production. *Phonology* 8, 73-112.
- Stemberger, Joseph Paul. 1993. Glottal transparency. *Phonology* 10, 107-138.
- Stemberger, J. P. and C. Stoel-Gammon. 1991. The underspecification of coronals: evidence from language acquisition and performance errors. In Paradis, C. and J-F. Prunet (eds.), *The Special Status of Coronals: Internal and External Evidence*. [Phonetics and Phonology, ed. S. R. Anderson and P. A. Keating, Vol. 2], 181-199. New York: Academic Press.
- Steriade, Donca. 1995. Underspecification and markedness. In Goldsmith, John A. (ed.) *The Handbook of Phonological Theory*, 114-174. Cambridge, Mass.: Basil Blackwell, Ltd.
- Yip, Moira. 1991. Coronals, consonant clusters, and the coda condition. In Paradis, C. and J-F. Prunet (eds.), 1991. *The Special Status of Coronals: Internal and External Evidence*. [Phonetics and Phonology, ed. S. R. Anderson and P. A. Keating, Vol. 2], 61-78. New York: Academic Press.

Exceptionality in Finnish

Andrew Dolbey
UC Berkeley/University of Helsinki

Ronald Sprouse
UC Berkeley

Introduction

In this paper we provide an analysis of variations in the shape of certain Finnish morphemes, in which a stem-final segment surfaces alternatively as a consonant or a vowel. Representative examples from nominal morphology are given in (1):

(1) (a) <u>'sky'</u>	(b) <u>'instance'</u>	(c) <u>'person'</u>	
taiva s -ta	tapa s -ta	i h mi s -tä	(part. sg.)
taiva a -n	tapa u k s e-n	i h mi s e-n	(gen. sg.)
taiva a -na	tapa u k s e-na	i h mi s e-nä	(ess. sg.)

The stem in (1a) provides an example of a stem-final alternating segment: in the partitive singular it is realized as a coronal fricative, whereas in the genitive and essive singular it is realized as a copy of the preceding vowel. Such stems contrast with those shown in (1b) and (1c); in the latter cases, the stem-final segment always surfaces as a consonant. In the analysis that follows, we show how the differences in behavior displayed by the final segments in the above stem types follow naturally from differences in their underlying segmental structure. These differences, in turn, interact with the grammar's constraint system to yield the results shown in (1). It will also be shown that the constraint system is not entirely uniform across the grammar; rather, differences in constraint rankings in a number of morphological contexts, together with the contrasts in underlying segmental structure, contribute to the observed variations in stem shape.

The alternations at issue belong to the core of Finnish inflectional morphology, an examination of which poses two general problems, summarized in (2).

- (2) a) Encoding the distinction between alternating and non-alternating segments;
- b) Accounting for the exceptional behavior of stems in certain grammatical environments

A systematic way of encoding the distinction between segments that alternate and segments that do not is needed to answer the first problem. We argue that underspecification and prespecification offer the best tools for deriving the contrast between alternating and non-alternating segments. The principle claim behind this proposal is that phonological alternations are the result of the interaction of segmental underspecification and constraint-driven feature fill-in.

Exceptional data are handled through prespecification, rather than diacritic rule features or lexeme-specific constraints; abstract, non-surfacing structure is eliminated when possible. A number of studies have shown that underspecification of this sort is necessary in order to obtain descriptive adequacy (Clements and Sezer 1982; Ringen 1988; Kiparsky 1993; Inkelas 1994; Inkelas, Orgun & Zoll 1994). However, in addition to providing descriptive adequacy, underspecification also has the desirable consequence of minimizing deletion: lexicon and grammar optimization (Prince & Smolensky 1993; Inkelas 1994) lead to representations which minimize deletion and maximize structure-building processes through grammar-sensitive archiphonemic representations.

The second problem posed by Finnish inflectional morphology involves a different kind of exceptionality: we must provide an explanation for the exceptional phonological behavior of stems in certain grammatical contexts. A number of works have shown that phonological patterning may vary systematically across grammatical constructions (Kiparsky 1982; Selkirk 1982; Selkirk 1984; Mohanan 1986; Cohn 1989; Inkelas & Orgun 1995; Itô & Mester 1995; Itô, Mester & Padgett 1995; Dolbey 1996). In order to account for this grammar-internal phonological variation, we draw on the insights of a number of works which make use of cophonologies, grammar-internal variations in the ranking of phonological constraints. In particular, we argue that the Finnish data provide further evidence for the theory of cophonologies presented in Inkelas and Orgun (1995). In this work, it is demonstrated that the order of phonological subgrammars is determined by morphological constructions, not by stipulated, extrinsic level-ordering. Evidence for this is revealed in the phenomenon of level non-ordering, in which the same cophonology is shared by non-contiguous morphemes. We will show below that Finnish inflectional morphology provides another case of level non-ordering, and hence is further evidence for the construction-driven theory of cophonologies argued for by Inkelas and Orgun.

Within the framework of generative grammar, the most comprehensive previous analysis of the data examined here is found in Keyser and Kiparsky (1984). Our analysis differs from theirs in matters of both specific detail and general theoretical orientation. In particular, whereas Keyser and Kiparsky's analysis is carried out using a rule-based, CV framework, our analysis is formulated within the framework of Optimality Theory (OT; P&S 1993; McCarthy and Prince 1993, 1995) and Moraic Theory (Hyman 1985, Zec 1988, Hayes 1989).

Morphophonology of Nominals

The data for the analysis are taken from standard reference grammars of Finnish (Karlsson 1983, Fromm 1982). In (3), we provide the segment inventory and a list of possible syllable types for standard Finnish.¹ Consonant and vowel length are

(6) Nouns with C~V stem alternations

	<u>'business'</u>	<u>'rich'</u>	<u>'spring'</u>
NOM SG	vene?	taivas	kevät
PART SG	venet-tä	taivas-ta	kevät-tä
GEN SG	venee-n	taivaa-n	kevää-n
ESSIV SG	venee-nä	taivaa-na	kevää-nä
INESS SG	venee-ssä	taivaa-ssa	kevää-ssä
ABLAT SG	venee-ltä	taivaa-lta	kevää-ltä

The important point to notice in these data is the variation in stem shape according to grammatical environment: in the nominative and partitive cases, the stem ends in a consonant; by contrast, in all other case forms (only four of them are given above) the stem ends in a lengthened form of the stem-final vowel. Note that in the first example, *vene?*, the stem-final consonant in the nominative is a glottal stop, while in the partitive it is a coronal stop. Glottal stops appear only word-finally in Finnish, and only when the following prosodic word begins with a vowel; otherwise, they surface as a copy of (i.e. forming a geminate with) the following word's initial consonant. The contrast is illustrated by the compounds given in (7):²

(7)	vene? ankkuri	vene? näyttely
	[vene?ankkuri]	[venenäyttely]
	'boat anchor'	'boat exhibition'

At first glance, one might expect the partitive to behave like other inflected forms, and attach to a lengthened final vowel; instead, the final segment of the stem is realized as a consonant. Therefore, whatever analysis is provided for the stem alternation, we must explain the unexpected behavior of the partitive as well. Note that the partitive case is morphologically and syntactically a member of the inflectional paradigm for nominals; it is not a special derivational suffix. Note also that the behavior of the partitive cannot be explained by making reference to its CV shape, as the essive case is also CV.³

There are several possible sources one could consider to account for the stem-alternations shown in (6). The least interesting analysis would be to posit listed stem allomorphs for these stems. This option is unappealing, however, in light of the fact that both productive morphology and loan-word incorporation can give rise to such alternating stems. Another option would be to posit a separate cophology specific to a particular subclass of lexemes. Although we present strong evidence for the existence of cophological variation in Finnish, we believe that the variation in stem shape demonstrated above should not be analyzed as the result of a cophology restricted to a class of lexemes, as there is

no independent property (semantic, syntactic, morphological, or phonological) grouping these lexemes together, to the exclusion of others. As argued by Inkelas, Orgun, and Zoll (1994), using cophologies for this kind of arbitrary grouping of lexemes would lead to an explosion of subgrammars. A third option would be to posit abstract allomorphs of the inflectional suffixes which trigger changes in stem shape. However, this would fail to explain the behavior of the non-alternating stems.

In the analysis presented below, we will seek a solution that locates the behavior in the stems themselves. That is, we argue that it is the underlying representation of the stems' segmental structure which gives rise to the alternations witnessed in (6). Deriving the alternations from the representation of the stems allows us to account for the complexities of the data as well as the lexical restrictions which the alternations are subject to. Given lexicon and grammar optimization, this is presumably the course the language learner follows as well.

Analysis: Underspecified root nodes

We propose that the behavior of alternating nominal stems requires an underlying form that distinguishes them from other stems. Specifically, these stems terminate in an underspecified segment that consists of an underlying root node and mora, with minimal featural specifications, as in (8):

(8)	a)	μ μ μ	b)	μ μ μ μ	c)	μ μ μ
UR		/vene /		/taiva •/		/kevã •/
				[+cont]		[-cont]

The stem *vene* in (8a) contains an unlinked mora that surfaces alternately as vowel length or a glottal stop. *Taiva* in (8b) ends in a [+cont] root node and mora, alternating between vowel length and [s]; *kevã* in (8c) ends with a [-cont] root node and mora, alternating between vowel length and [t]. These are the alternations shown in (6).

These representations contain abstract structure that fails to surface in certain environments; however, we argue that these abstractions are necessary to appropriately distinguish alternating stems from non-alternating stems. If these stem-final segments were represented as full segments in UR incorrect predictions would result – it would be impossible to distinguish alternating /s/ and /t/ from their non-alternating counterparts, shown in (9):

(9)	<u>root</u>	<u>genitive</u>	<u>partitive</u>	<u>essive</u>
	hyvyyt-	hyvyyde-n	hyvyyt-tä	hyvyyte-nä
	ihmis-	ihmise-n	ihmis-tä	ihmise-nä
	tapaus	tapaukse-n	tapaus-ta	tapaukse-na

The stems in (9) end with the fully-specified consonantal segments indicated by underlining. These stem-final consonants do not alternate with vowel-length as a result of affixation. They do alternate, most notably in *hyvyyden*, as a result of consonant gradation, but they never lose their consonantal quality. Before C-initial suffixes, except for the partitive, which we will analyze later, an epenthetic [-e-] is inserted, and the stem-final consonants surface as onsets. It should be noted that no mora corresponds to these consonants when they surface in the onset; therefore, these segments are not specified as moraic in UR.

Alternative representations of the stem-final alternating segments might include less structure than what is proposed in (8). For example, these segments might consist of nothing more than a root node, a mora, or just a few floating features. However, root nodes and/or moras are not sufficient on their own to distinguish one alternating segment from another. [ʔ], [s] and [t] are not predictable from the phonological environment. Similarly the presence of floating features alone in UR is insufficient to explain the presence of vowel length when the features fail to surface. As a result the most transparent representation of these segments includes all three types of structure – a mora, root node and the feature [cont]. The aim of these underspecified representations is in the spirit of Kiparsky (1993) in the attempt to minimize deletion wherever the data allows it. Furthermore, our goal is to use underspecification to describe alternations, as in Inkelas (1994), not to minimize lexical representation or to derive a theory of markedness.

With these lexical representations in mind, it is possible to explore the cophologies that select the surface realizations of the different morphological constructions.

Constraints that derive the surface realization: Nominative forms and Word phonology

We consider first the nominative construction, where the underspecified segments surface as consonants:

(10)	Nominative:	vene?	taivas	kevät
------	-------------	-------	--------	-------

Three constraints are sufficient to select the surface forms in (10) based on the lexical representations in (8). The first of these is a Faithfulness constraint, MAX (M&P 1995), which disallows deletion of root nodes and moras present in UR

from surface forms. The other two constraints are well-formedness constraints. The first is NOCODA (M&P 1993; P&S 1993), which disallows coda consonants, and the second is NOLONGV, a constraint prohibiting long vowels (Sherer 1994). We refer to the nominative's cophonology as the Word level cophonology.

Operative Word level constraints:

- (11) MAX(μ ,RT) Don't delete segments. (M&P 1995)
 (12) NOLONGV Long vowels are prohibited. (Sherer 1994)
 (13) NOCODA Syllables should be open. (M&P 1993; P&S 1993)

The tableau in (14) demonstrates the ranking order of these three constraints. MAX dominates NOCODA, and for this reason, candidate (b) is favored over (c), which fails to realize the underspecified root node and mora present in UR. Candidate (a) is discarded since it also violates MAX. Consequently, candidate (b) is the winner since it violates only the low-ranked NOCODA. Although the long vowel in (a) is indicated by two orthographic [a]s, this is merely to make the display more readable, and we assume that vowel lengthening is the result of deletion of the final root node with spreading of the preceding vowel's features onto the stray mora. This is a non-crucial assumption since the alternative representation of the long vowel with two root nodes (and hence no MAX violation; see Selkirk 1990 for a proposal to represent long segments with two root nodes) also fails because of the ranking of NOLONGV above NOCODA. Further evidence for this ranking will be seen in the discussion of compounding.

Nominative

Max(μ ,Rt) >> NoLongV >> NOCODA = no deletion, no compensatory lengthening

(14)

	$\mu\mu$ $\mu\mu$ ta iva * [+cont]	MAX(μ ,RT)	NOLONGV	NOCODA
(a)	taiva <u>a</u>	*!		
(b)	\Rightarrow taiva <u>s</u>			
(c)	taiva <u>_</u>	*!		

The [cor] specification in the winning candidate *taivas* is not present in UR. Throughout this paper we follow without formalizing feature-filling conventions of the type found in Prince & Smolensky (1993), Smolensky (1993) or Kiparsky (1994), based on markedness constraint hierarchies (see also Paradis & Prunet 1991 for the unmarked status of [cor]). For the segments marked [cont] in UR, [cor] is the default specification for Place.⁴ The default features assigned to alternating segments are summarized in (15).

(15) Feature-filling insertion	<u>features inserted</u>
taiva •[+cont] → taiva _s	[cor]
kevä •[-cont] → kevä _t	[cor]

The nominal compound construction also constricts the Word level cophonology. In this construction alternating /s/ and /t/ surface as [s] and [t] even when followed by a C-initial word. A complication arises with regard to the ‘glottal-stop-final’ stems, however. In this environment [ʔ] fails to surface, and instead the initial consonant of the following word geminates. The representation of these stems in UR and the Word level cophonology motivated in (14) predicts this gemination. As the tableau in (15) illustrates, gemination violates only the low-ranked NOCODA constraint, as in the winning candidate (c). Deletion of the bare mora in (b) rules out that candidate, and vowel-lengthening in (a) is disfavored by NOLONGV.

Compound -- C-initial second PwD

Max(μ) = compensatory lengthening with empty mora

NoLongV >> NoCoda = gemination, not vowel lengthening

(15)	$\mu \mu \mu$ vene , mies	MAX(μ ,RT)	NOLONGV	NOCODA
(a)	ven <u>e</u> mies		*!	
(b)	vene_ <u>m</u> ies	*!		
(c)	⇒ ven <u>e</u> m <u>m</u> ies			

In summary, the Word level cophonology so far developed is associated with the Nominative and Compounding constructions. The high ranking of MAX in this cophonology favors the retention of all the underlying structure of the underspecified segments. Furthermore, the ranking of NOLONGV above NOCODA favors gemination of the consonant following a ‘glottal-stop-final’ stem in compounds.

(16) Word level cophonology (1) – associated with Nominative and Compounds

MAX undominated = no deletion of alternating segments

NOLONGV >> NOCODA = gemination, not vowel-lengthening

We turn now to the Stem level cophonology, which is associated with all the inflectional affixes except the partitive.

Stem Phonology

In all the inflectional cases except the nominative and the partitive, stem-final alternating segments surface as vowel length rather than as consonants. Our analysis seeks to explain three questions regarding these facts: (i) Why do the underspecified segments fail to surface? (ii) Why is there compensatory lengthening of the stem-final vowel? (iii) Why is the underlying mora filled in with the features of the stem-final vowel, rather than the following consonant? In other words, why is *taiva* 'sky', which is [taivas] in the nominative form, [taivaana] in the essive singular case, not [*taivasna] or [*taivanna]? The answer to these questions lies in the ranking of NOCODA and NOLONGV with respect to each other and with respect to the Faithfulness constraints in the Stem cophonology.

The essive singular case serves to illustrate the behavior of these stems when a -CV affix is adjoined to them. As the tableau in (17) shows, MAX μ is undominated in the Stem cophonology, as it is in the Word cophonology, a fact which requires the stem-final mora to be present in the output. Candidate (d) is out for deleting that mora. MAXRT, however, is not undominated in the Stem cophonology, allowing deletion of the stem-final root node. The only candidate that shares a stem-final [s] with the nominative [taivas], candidate (c), satisfies MAXRT, but it violates the higher-ranked NOCODA and is also ruled out. The two remaining candidates preserve the underlying mora by lengthening a segment adjacent to the deleted root node. Candidate (a) geminates the following consonant, but this incurs a violation of NOCODA. Consequently (b), with vowel lengthening, is the winner. Again, the decision to represent long segments as a single root node linked to two moras is not a crucial one. A two root node representation of these segments would simply mean that (a) and (b) do not violate MAXRT but (d) still would. The ranking of NOCODA above NOLONGV would still be the important one to distinguish these two candidates.

Essive

MAX μ , NOCODA >> NOLONGV, MAXRT = deletion and V-lengthening

(17)

	$\mu\mu\ \mu\ \mu$ ta iva •[+cont] + nA	MAX μ	NOCODA	NOLONGV	MaxRt
(a)	taiva <u>anna</u>		*!		
(b)	⇒ taiva <u>a</u> na				
(c)	taiva <u>s</u> na		*!		
(d)	taiva <u>_</u> na	*!			

One other candidate remains to be considered. NOCODA is crucial in ruling out candidate (c) in (17). It would be possible to insert a vowel between the stem-final

segment and the affix to prevent the NOCODA violation. Such a candidate is ruled out, however, by another Faithfulness constraint, DEP, which disfavors epenthesis.

(18) DEPRT Don't insert a root node. (M&P 1995)

The tableau in (19) demonstrates that ranking DEP above NOLONGV rules out epenthesis as a possible repair for NOCODA violations when the stem ends in an alternating consonant.

Essive

DEPRT >> NOLONGV = no epenthesis to save NOCODA violation

(19)

	μ μ μ μ ta iva •[+cont] + nA	Maxμ NOCODA	DEPRT	NOLONGV MAXRT
(a)	taiva <u>s</u> ena		*!	
(b)	⇒ taiva <u>a</u> na			*

When the inflectional affix is of the form -C instead of -CV, then the surface form will always have a NOCODA violation triggered by the affix, as in the genitive form [taivaan]. In this case no extra NOCODA violations are incurred by allowing the underlying root node to surface. Since NOCODA is crucial to rule out the candidate that preserves the underlying root node in (17), we need to explain why the underlying root node of *taiva* surfaces as vowel length rather than as a consonant. The answer is that [sn] is an ill-formed coda, a fact that is expressed by the *COMPLEX constraint.

(20) *COMPLEX no complex coda (Prince & Smolensky 1993)

Genitive

*COMPLEX = complex codas not allowed

(21)

	μ μ μ μ ta iva •[+cont] + n	*COMPLEX NOCODA	DEPRT	NOLONGV
(a)	taiva <u>s</u> en	*	*!	
(b)	taiva <u>s</u> n	*!		
(c)	⇒ taiva <u>a</u> n			*

In (21) candidate (b) is ruled out for violating *COMPLEX. DEPRT again rules out the candidate with epenthesis (a). The winner therefore is (c), with deletion of the underlying root node and lengthening of the vowel onto the empty mora.

The crucial difference between alternating consonants and full consonants that never alternate with vowel length is the fact that full consonants are fully specified with Place features in UR. This is the key to understanding the difference in their behavior. The constraint ranking of the Stem level cophonology as it is developed so far appears to predict that all coda consonants should delete since NOCODA outranks MAXRT. Another constraint is necessary to prevent deletion of fully-specified consonants, such as the final segment of *taivan*, the winner in (21). Deletion of fully-specified segments requires deletion of Place features as well as the root node, and this is not allowed. The relevant constraint governing this type of Faithfulness is MAXPL.

(22) MAXPL Don't delete Place features. (Zoll 1996; cf. M&P 1995)

The tableau in (23) illustrates the behavior of fully-specified stem-final consonants. Candidates (c-e) delete the final stem segment in order to avoid NOCODA violations. All three of these candidates are ruled out, however, because they violate the high-ranked MAXPL constraint. Of the two remaining candidates, (b) has a DEPRT violation, resulting in one fewer NOCODA violations than (a). Because of the ranking of NOCODA over DEPRT, (b) is the winner. The result is that epenthesis is favored when deletion of the stem-final consonant is not allowed. This contrasts with the behavior of stems ending in alternating segments. MAXPL is irrelevant for these stems, so the ranking of DEPRT over MAXRT ensures that deletion is favored over epenthesis to avoid possible NOCODA violations.⁵

Fully-specified segment

MAXPL >> NOCODA = no deletion

NOCODA >> DEPRT = epenthesis when segment not deleted

(23)

	$\mu\mu\mu$ $ $ samal + nA [Pl(Cor)]	MAX(PL)	NOCODA	DEPRT	NO LONGV	MAXRT
(a)	sammalna		**!			
(b)	⇒ sammalena		*	*		
(c)	sammanna	*!(Cor)				
(d)	sammaana	*!(Cor)				
(e)	sammana	*!(Cor)				

In summary the Stem level cophonology is associated with most inflectional morphology and differs from the Word level cophonology in the ranking of several constraints. The ranking of MAXPL over NOCODA prevents deletion of

fully-specified consonants in codas. Ranking NOCODA over NOLONGV favors vowel-lengthening over gemination when an alternating segment is deleted, and ranking NOCODA above DEPRT requires epenthesis after fully-specified stem-final consonants.

- (24) Stem level cophonology (1) – associated with Nominative and Compounds
 MAX(PL) >> NOCODA = deletion of underspecified segments only
 NOCODA >> NOLONGV = vowel-lengthening, not gemination
 NOCODA >> DEPRT = epenthesis after fully-specified stem-final Cs

We consider next the exceptional behavior of the partitive affix and account for this behavior using the analysis that has been developed so far.

Phonology of the Partitive

The partitive affix constricts the Word level cophonology, in contrast with all other inflectional affixes. In doing this the partitive patterns with other Word level morphological operations, including compounding and cliticization. The main diagnostics of Word level behavior are illustrated in (25-27):

- (25) Phrasal phonology
- | | | |
|--------------|----------------|---------------|
| vene täällä | taivas tietää | kevät toimi |
| [venetäällä] | [taivastietää] | [kevättöimi] |
| 'boat here' | 'heaven knows' | 'spring work' |
- (26) Clitics
- | | | |
|-----------|------------|------------|
| vene-kö | taivas-ko | kevät-kö |
| [venekkö] | [taivasko] | [kevätkkö] |
| boat-Q | sky-Q | spring-Q |
- (27) Compounds
- | | | |
|--------------|-------------------|--------------|
| vene-talas | tehdas-työläinen | kevät-talvi |
| [venettalas] | [tehdastyöläinen] | [kevättalvi] |
| boat shed | factory worker | late winter |

The first diagnostic is illustrated by the glottal-final stems. Whenever these stems are followed by a consonant in the Word level cophonology, the following consonant geminates. This behavior contrasts with the Stem level cophonology in which gemination is not allowed and vowel length surfaces instead. The second diagnostic is illustrated by the stems ending in alternating /s/ and /t/. Epenthesis never occurs between one of these stems and a following consonant at the Word level. Again, this contrasts with the Stem level phonology, where epenthesis

always occurs between a fully-specified stem-final consonant and a following consonant.

Because the partitive patterns with Word level constructions in all of these diagnostics, we conclude that it is associated with Word level phonology. After glottal-final stems, the initial consonant of the partitive is a geminate, as in *venettä*. After stems ending in alternating /s/ or /t/, the alternating segments surface as consonants and no epenthesis occurs between the stem and the partitive affix, as in *taivasta*.⁶

The tableaux in (28) and (29) show what happens when the partitive is added to a glottal-final stem and C-final stem. The ranking of MAX μ and NO LONGV in the Word phonology were motivated in the discussion of the nominative and compounds. In (28) the winning candidate geminates the initial consonant of the partitive, violating only low-ranked NOCODA. Candidate (b) is ruled out for violating NO LONGV, and (c) fails for deleting the underlying stem-final mora.

Partitive -- Word-level constraint ranking

DepRt >> NoCoda = no epenthesis

NoLongV >> NoCoda = gemination, not vowel lengthening

(28)

	$\mu \mu \mu$ vene + tA	MAX(μ)	DEPRT	NO LONGV	NOCODA
(a)	\Rightarrow venet $_$ tä				*
(b)	venee $_$ tä			*!	
(c)	vene $_$ -tä	*!			

(29)

	$\mu \mu \mu$ samal + tA [Pl(Cor)]	MAX(μ)	DEPRT	NO LONGV	NOCODA
(a)	\Rightarrow sammalta				**
(b)	sammaleta		*!		

Tableau (29) shows the partitive affixed to a C-final stem and is important for establishing the relative ranking of DEPRT in the Word level cophonology. Unlike the Stem level cophonology, epenthesis is not allowed between stem and affix. Ranking DEPRT above NOCODA assures this outcome. Candidate (b) violates DEPRT and is ruled out. The winning candidate (a) has no violation of DEPRT, although it does have an additional NOCODA violation.

In summary, the exceptional behavior of the partitive affix motivates its inclusion in the Word level cophonology, a move which simultaneously predicts the gemination of its initial segment after glottal-final stems and the lack of epenthesis after C-final stems.

(30) Word level cophonology (2) – associated with nominative, compounds and partitive

MAX undominated	= no deletion of alternating segments
NO LONGV >> NO CODA	= gemination, not vowel-lengthening
DEPRT >> NO CODA	= no epenthesis

Level non-ordering

Recent work on cophonological variation (Mohanani 1986; Inkelas and Orgun 1995, Orgun 1997) has brought attention to the phenomenon of level non-ordering, in which morphemes sharing the same cophonology, or ‘level’, do not cluster together syntagmatically. The principle insight gained from this work is that the ordering of cophonologies is a function of the morphological constructions in which they participate, not some extrinsic constraint on cophonologies; hence level ordering, although possible, is not a necessary result of cophonological variation. Here we show that Finnish provides another case of level non-ordering, and consequently offers further evidence for a construction-based analysis of cophonology interaction.

The data at issue are the alternating stems discussed above, in combination with possessive suffixes and clitics. Kanerva (1987) has provided strong evidence for the status of possessive suffixes as *inflectional affixes*, rather than clitics. Among the principle arguments for their status as affixes are the following two facts: 1) possessive suffixes, like inflectional suffixes, include cases of allomorphic variation, whereas clitics do not; and 2) possessive suffixes attach to stem forms in the nominative, whereas clitics attach to the normal nominative word form. These two contrasts are illustrated in (31):

(31)	<u>3rd Px: -Vn ~ -nsA</u>	<u>Emphatic clitic: -pA</u>
	talo=nsa ‘his/her house’	talo=pa ‘house indeed’
	talo-ssa=an ‘in his/her house’	talo-ssa=pa ‘in the house indeed’
	hevose=nsa ‘his/her horse’	hevonen=pa ‘horse indeed’

Interestingly, when possessive suffixes are attached to stems with alternating final segments, the stem-final segment surfaces as a vowel; that is, possessive suffixes conscript the stem level cophonology. By contrast, clitics attach to stems whose final segments surface as a consonant; hence, clitics conscript the word level cophonology. The contrast is illustrated in (32):

(32)	Px:	venee-ni	taivaa-ni	kevää-ni	
		NOM 1SGPx	NOM 1SGPx	NOM 1SGPx	
		venee-si	taivaa-si	kevää-si	
		NOM 2SGPx	NOM 2SGPx	NOM 2SGPx	
	Cl:	venek-kö	taivas-ko	kevät-kö	
		NOM-Q	NOM-Q	NOM-Q	(Q = interrogative clitic)
		venep-pä	taivas-pa	kevät-pä	
		NOM '!!'	NOM '!!'	NOM '!!'	(!! = emphatic clitic)

Cases where a partitive word form surfaces with both a possessive suffix and a clitic illustrate clearly the phenomenon of level non-ordering. Morphotactic constraints require the morpheme order shown in (33):

(33) Morpheme order: Stem – Case Affix – Possessive Suffix – Clitic

This morpheme ordering, however, results in a word form with non-contiguous, layered cophologies, as shown in (34):

(34)	Layered cophologies:	venet-tä-än-kö
		root PT Px Cl
		St Wd St Wd

Conclusion

We began the paper by raising the following two questions:

- a) how do we encode the distinction between alternating and non-alternating segments?
- b) how can we explain the exceptional behavior of stems in particular grammatical contexts, namely when followed by the partitive?

It was shown that the answer to the first question requires positing an underspecified segment consisting of a root node and a mora. This underspecified segment interacts with the phonological constraints provided by the grammar to produce the variations in stem shape obtained in the alternating stems. The behavior of non-alternating stems was accounted for by positing fully specified stem-final segments; high-ranking prohibitions against deletion of features results in non-alternating stems. Although underspecification provides the key to answering the first question, it does not help us with the second. For this it was shown that cophologies are necessary. The cophology associated with a given morphological construction will determine whether the stem surfaces with vowel-lengthening (Stem level phonology) or gemination (Word level phonology). This construction-driven understanding of cophology interaction also helps to explain the possibility of level non-ordering, illustrated in the previous section. Together, underspecification and cophologies provide all and only the tools needed to deal with alternations and exceptionality.

References

- Clements, George. 1976. Neutral vowels in Hungarian vowel harmony. NELS 7.
- Clements, George and Samuel Keyser. 1983. *CV Phonology*. Cambridge, Mass.: MIT Press.
- Clements, George N. and Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. In H. v. d. Hulst and N. Smith (eds.), *The structure of phonological representations. Part II*. Dordrecht: Foris. 213-255.
- Cohn, Abigail. 1989. Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7. 167-216.
- Dolbey, Andrew. 1996. Output optimization and cyclic allomorph selection. To appear in WCCFL 15.
- Fromm, Hans. 1982. *Finnische Grammatik*. Heidelberg: C. Winter, Universitätsverlag.
- Harris, James. 1977. Spanish vowel alternations, diacritic features, and the structure of the lexicon.
- Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20:253-306.
- Hyman, Larry. 1985. *A theory of phonological weight*. Dordrecht: Foris.
- Inkelas, Sharon. 1994. The consequences of Optimization for Underspecification. NELS 25.
- Inkelas, Sharon and Young-mee Cho. 1993. Inalterability as prespecification. *Language* 69.
- Inkelas, Sharon and Cemil Orhan Orgun. 1995. Level ordering and economy in the lexical phonology of Turkish. *Language*.
- Inkelas, Sharon, Cemil Orhan Orgun, and Cheryl Zoll. 1994. Exceptions and static phonological patterns: cophonologies vs. prespecification. In I. Roca, ed. *Constraints and Derivations in Phonology*.
- Itô, Junko. 1986. Syllable theory in prosodic phonology. PhD dissertation. University of Massachusetts, Amherst. GLSA.
- Itô, Junko and Armin Mester. 1995. Japanese Phonology. In J. Goldsmith ed. *The Handbook of Phonological Theory*.
- Itô, Junko, Armin Mester, and Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26.4. 571-613.
- Kager, Rene. 1992. Shapes of the generalized trochee. WCCFL 11.
- Kanerva, Jonni. 1987. Morphological Integrity and syntax: The evidence from Finnish possessive suffixes. *Language* 63(3).
- Karlssohn, Fred. 1983. *Finnish Grammar*. Translated by Andrew Chesterman. Porvoo [Finland]: Werner Soderstrom Osakeyhtio.
- Keyser, Samuel J. and Paul Kiparsky. 1984. Syllable structure in Finnish phonology. In Mark Aronoff and Richard T. Oehrle, eds. *Language Sound Structure*.
- Kiparsky, Paul. 1982. Lexical Phonology and Morphology. In *Linguistics in the Morning Calm* 1, ed. I.-S. Yang, 3-91. Seoul, Korea: Hanshin
- Kiparsky, Paul. 1985. Some Consequences of Lexical Phonology. *Phonology Yearbook* 2: 82-138.
- Kiparsky, Paul. 1993. Blocking in non-derived environments. In Hargus, Sharon and Ellen Kaisse, eds. *Studies in Lexical Phonology: Phonetics and Phonology* 4. San Diego: Academic Press.
- Kiparsky, Paul. 1994. Remarks on Markedness. Handout from talk presented at *Trilateral Phonology Weekend*, University of California, Santa Cruz.
- McCarthy, John and Alan Prince. 1993. *Prosodic morphology I: Constraint interaction and satisfaction*. Ms. University of Massachusetts, Amherst and Rutgers University. [To appear, MIT Press]
- McCarthy, John J. and Prince, Alan. 1995. Faithfulness and Reduplicative Identity. In Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (ed.). *Papers in Optimality Theory*. UMass, Amherst: GLSA. 249-384.

- Mohanan, K. P. 1986. *The theory of Lexical Phonology*. Dordrecht: Kluwer.
- Orgun, Orhan. 1997. *Sign-Based Morphology and Phonology*. Ph.D. Dissertation. University of California at Berkeley.
- Paradis, Carole and Jean-Francois Prunet. 1991. *The Special status of coronals*. San Diego: Academic Press.
- Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Ms. Rutgers University and University of Colorado, Boulder. [To appear, MIT Press]
- Ringen, Catherine. 1988. Transparency in Hungarian vowel harmony. *Phonology* 5:327-342.
- Selkirk, Elisabeth. 1982. *The syntax of words*. Cambridge, MA: MIT Press.
- Selkirk, Elisabeth. 1984. *Phonology and syntax*. Cambridge, MA: MIT Press.
- Selkirk, Elisabeth. 1990. A two root theory of length. *University of Massachusetts Occasional Papers* 14. 123-171.
- Sherer, Tim. 1994. *Prosodic phonotactics*. PhD dissertation. University of Massachusetts, Amherst. GLSA.
- Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. ROW 1.
- Spencer, Andrew. 1994. Syllabification in Chukchee. In R. Weise, ed. *Theorie des Lexicons*.
- Sprouse, Ronald. 1996. Vowels that borrow moras: Geminate and weight in OT. NELS 26.
- Zec, Draga. 1988. *Sonority constraints on prosodic structure*. PhD dissertation. Stanford University.
- Zoll, Cheryl (1996). Parsing Below the Segment in a Constraint-Based Framework. PhD dissertation. University of California at Berkeley.

¹ For purposes of this paper, we ignore segment types such as voiced stops [b, d, g] and the voiceless velar fricative [f]. These have been more recently incorporated into the segment inventory; their status as 'genuine' Finnish segments is less clear. This omission has no bearing on the analysis proposed.

² The phonetic realization of the glottal stop varies in different dialects and has been described as a true glottal stop, a 'catch' or simply silence (Fromm 1982). It is never realized as vowel length. Although we represent this segment with the symbol ? and refer to it as a 'glottal stop', it is represented formally in our analysis as an empty mora.

³ Nor can the special behavior of the partitive be attributed to its stop-vowel shape, as a number of other affixes with nasal-vowel shape (identical to that of the essive) behave in the same way as the partitive. These affixes include verbal suffixes such as the perfect /-nut/ and the potential /-ne/.

⁴ The 'glottal-final' stems end in an empty mora. With the exception of one morphological environment, the features linked to this mora are provided by an adjoining segment. When this mora is word-final, as in the nominative, no features are filled in. We propose that a constraint against word-final long vowels prevents the lengthening of the vowel and that NOCODA and/or a constraint against inserting root nodes prevents the mora from being realized as a full consonant.

⁵ The epenthesis facts are slightly more constrained than our ranking suggests. Notice that /sammal + n/ → [sammalen], not [*sammalene]. Descriptively, epenthesis is allowed only between the root and the affix. We propose that constraints such as Contiguity and Anchor can predict these facts (cf. Spencer 1994).

⁶ Our analysis does not account for the allomorphy of the partitive after short vowels since it is orthogonal to the other issues in this paper. We believe that the two forms of the partitive are best analyzed as listed allomorphs that subcategorize for roots of a particular phonological shape.

RESTRICTING CONJUNCTION TO CONSTRAINT FAMILIES

Haruka Fukazawa and Viola Miglio
University of Maryland at College Park

1. Introduction

Within the OT framework (Prince and Smolensky 1993), a different ranking of constraints accounts for the different grammars in the world's languages. There are, however, some phonological data that cannot be explained by the ranking of single constraints, such as stress assignment in Diyari (Hewitt & Crowhurst 1995), vowel length phenomena in the Wellagga dialect of Oromo (Alderete 1996), vowel raising in Nzebi (Kirchner 1996), and Southern Palestinian Arabic RTR phenomena (McCarthy 1996). In such cases, the analyses of the data are made possible only by introducing Local Conjunction.

Local conjunction is defined as a combination of two single lower-ranked constraints that forces a violation of a higher one (Smolensky 1993, 1995)¹. If constraint A and constraint B are each ranked lower than constraint C, a candidate can violate either of them so as to satisfy C, since

$$(1) \quad C \gg A, B$$

However, if a candidate violates *both* A and B, the conjunction of these two violations may result in a fatal violation. The conjunction may, in fact, even force a violation of the higher ranked constraint C:

$$(2) \quad A\&B \gg C \gg A, B$$

(2) indicates that A and B are each separately violable so as to satisfy the higher ranked constraint C; however, violating both of them simultaneously is worse than violating C.

Although several studies are based on local conjunction, its scope and definition are still ambiguous. Especially, if local conjunction is a type of constraint, it must be in UG. However, if it is in UG, it must be cross-linguistically valid. A question arises: are all local conjunctions in UG? If so, UG would grow extremely large.

This paper proposes that the possibility of local conjunction is in UG, i.e. an "&" operator for conjunction. However, the choice of the two constraints to conjoin is language-specific.

This proposal takes the burden off a larger UG and seems to be corroborated by the cross-linguistic rarity of each particular type of local conjunction. Because of the nature of local conjunction, as the union of two lower-ranked constraints overriding hierarchically higher-ranked ones, it should be considered as a last resort operation. In other words, local conjunction should come into play only when every ranking of single constraint fails to explain the data in a language.

However, it seems necessary to restrict local conjunction even further. If any constraint can be conjoined with any other, even the language-specific grammar becomes extremely unrestricted. Smolensky (1993, 1995) has pointed out one important restriction of conjunction: locality. The two constraints to be conjoined

must be violated in the same location, since constraint interaction is stronger locally than non-locally.

McCarthy suggests (1996) that the two constraints to be conjoined must be phonetically conjoinable. The constraint conjunction that he uses in his analysis of Southern Palestinian Arabic is *RTR [HI] & *RTR [FRONT]. According to him, RTR (retracted tongue-root) is phonologically one of the distinctive features for "emphasis" and phonetically a kind of uvularization. Hence, *RTR [HI] indicates that the distinctive feature RTR does not coexist with the feature HIGH in the same segment. *RTR [FRONT] means that the two features RTR and FRONT are mutually exclusive within a segment. McCarthy states that the two constraints, *RTR [HI] and *RTR [FRONT] are phonetically conjoinable, since both of them are a formalization of the fact that it is not possible "to constrict the pharynx when the tongue body is being pulled in the wrong direction" (*ibid.*, p.7). Thus, McCarthy's conjunction is restricted to phonetically conjoinable constraints.

It might be true that the restriction of phonetic conjoinability is valid for some local conjunctions; however, it cannot restrict all kinds of conjunctions, since not all constraints are phonetically motivated. For instance, it is not possible to define OCP[place] from a phonetic viewpoint, since this constraint is not phonetically grounded. A conjunction of [OCP] constraints, therefore, would not be restricted by this notion.

Thus, other general restrictions on the conjoinability of two constraints should be explored. The purpose of this paper is to propose a strict restriction on the conjoinability of constraints in local conjunction, namely that the constraints to be conjoined must belong to the same constraint family. We are basing our proposal on the analysis of two new sets of data from different language families, as well as on several analyses of local conjunction proposed by other researchers.

The claim is that, whenever data are analyzed by means of local conjunction, the following points should be considered:

1. Motivation: every alternative ranking of single constraints fails to produce the correct analysis;
2. Restrictions: (a) locality must be respected;
 (b) phonetic conjoinability may be taken into consideration in some cases;
 (c) the two constraints to be conjoined must belong to the SAME CONSTRAINT FAMILY.

In order to clarify these points, this paper will introduce the analysis of two phonological phenomena from two different languages: spirantization in Yucatec Maya (Fukazawa 1996) and vowel raising in the Northern Mantuan dialect of Italian (Miglio 1996). In section 2, the analysis of spirantization in Yucatec Maya consonant clusters will be examined; section 3 will present an analysis of front vowel raising in Mantuan, and in the final section, the summary of previous research on local conjunction will clarify how the results obtained by other authors also support our proposal.

2. Yucatec Maya

2.1 Phonological Alternations in Consonant Clusters

In Yucatec Maya, when a stop and an affricate are followed by a homorganic stop (or affricate), they become a pharyngeal fricative and a

homorganic fricative, respectively. Those alternations are observed neither when they are followed by a non-homorganic stop (or affricate), nor when they precede a fricative.

The following data show the phonological alternations observed in the language:

(3) **Yucatec Maya** (Straight 1976):

- a. taarj k pak'ik k kool → taarj k pak'ik h kool
 "we're planting our clearing."
 b. tun kolih k'aaš → tun kolih h'aaš "he's clearing bush"
 c. le? iŋ w oŋ čo → le? iŋ w oŋ h o "that house of mine/my house there"
 d. ?uc t iŋ w ič → ?us t iŋ w ič
 "I like it (lit., "goodness is at my eye.)."
 e. ?u k'áat u karj kaasteyàanoh → ?u k'áat u karj kaateyàanoh
 "He wants to learn Spanish."

In the above data, "homorganic" refers only to major articulation. The coronal obstruents count as homorganic regardless of their value of anterior. It is also irrelevant whether the consonants differ in glottalization (k or k').

The sequences in which the alternations are observed are formalized as follows:

- (4) a stop + a homorganic stop: (5) an affricate + a homorganic stop
- | | | | |
|--------|--------|-----------------|--------|
| [dor] | [dor] | [cor] | [cor] |
| | | | |
| Root • | Root • | Root • — [cont] | Root • |
| | | | |
| [stop] | [stop] | [stop] | [stop] |

As (4) and (5) show, there are two identical [stop] features and two identical [place] features adjacent in the sequence of a stop and a homorganic stop, as well as in a sequence of an affricate followed by a homorganic stop.

On the other hand, the alternations are not observed in the following sequences:

- (6) a fricative + a homorganic stop: (7) a stop + a non-homorganic stop
- | | | | |
|--------|--------|--------|--------|
| [cor] | [cor] | [dor] | [lab] |
| | | | |
| Root • | Root • | Root • | Root • |
| | | | |
| [cont] | [stop] | [stop] | [stop] |

In the sequence of a fricative and a homorganic stop (6), the two adjacent segments in the sequence have the same place features and two different manner features. In the sequence of a stop and a non-homorganic stop in (7), the two segments in the sequence have the same manner features and different place features.

Lombardi (1990a, b) analyzes the alternation as the result of delinking one of the [stop] features due to the effect of the OCP (Leben 1973, Goldsmith 1976, Mester 1986, McCarthy 1986). The next section will analyze these phonological phenomena within the OT framework, showing the necessity for local conjunction.

2.2 An Analysis with Single Constraints

First, the analysis based on the ranking of single constraints will be presented and it will be shown why it does not work well. Next, the analysis with the conjunction will be introduced to clarify why the conjunction is necessary for the Yucatec data.

Two kinds of OCP effects should be considered in the analysis of Yucatec Maya: OCP on [stop] and OCP on [place] features. The present proposal hinges on the fact that two adjacent segments with the same place features are affected by the OCP effect on place feature.

Moreover, the deletion of the [stop] feature is observed as the result of the OCP effects in Yucatec Maya (Lombardi 1990a, b). Therefore, at least three kinds of constraints interact here: OCP[place], OCP[stop] and IDENT[stop].

Since the effects of the OCP can be seen at work, they are assumed to be relatively high-ranked constraints in this language, and they must be satisfied at the expense of the violation of some lower-ranked constraint(s). Since one of the [stop] feature deletes, it is assumed that the lower-ranked constraint to be violated is a featural faithfulness constraint for [stop], namely, IDENT[stop].

Another featural faithfulness constraint for [place] should also be considered, IDENT[place], although a further argument would be necessary to explain why the place feature changes². Consequently, the following four constraints are necessary in the analysis:

OCP: Two adjacent identical elements are prohibited (Leben 1973, Goldsmith 1976, Mester 1986, McCarthy 1986).

(A). OCP[place]³: Two adjacent identical place features are prohibited:

(B). OCP[stop]: Two adjacent identical stop features are prohibited:

IDENT [F]: Correspondent segments have identical values for the feature [F] (McCarthy & Prince 1995).

(A). IDENT[place]³: Correspondent segments have identical values for the feature [place]:

(B). IDENT[stop]: Correspondent segments have identical values for the feature [stop]:

Thus, we assume the following ranking of those four constraints from what has been discussed so far:

(8) OCP[place], OCP[stop] >> IDENT[place], IDENT[stop]

Let us analyze actual data with this ranking. As the data in (3) show, when a stop is followed by a homorganic stop, the first stop becomes [h]. For example, /k kool/ becomes [h kool] (3a). These data are examined in the following tableau:

(9) A stop followed by a homorganic stop.

/k kool/	OCP[place]	OCP[stop]	IDENT[place]	IDENT[stop]
a. [☞] h kool			*	*
b. [☹] k kool	*!	*!		

Candidate (b), in which no alternation is observed, violates both of the two higher-ranked constraints, namely, OCP[place] and OCP[stop]. Since both of them are high-ranked, the violation of only one of them is enough for the candidate to lose. On the other hand, candidate (a), in which the alternation is observed, violates neither OCP[place] nor OCP[stop], and therefore it wins. The ranking in (8) correctly provides the optimal candidate so far. From tableau (9), we should conclude that at least one of the OCPs must outrank IDENT[place] and IDENT[stop] to account for the correct output.

Let us look at the next data: a stop and a non-homorganic stop. In this sequence, no phonological alternation is observed.

(10) Wrong result: A stop and a non-homorganic stop.

/k pak'ik/	OCP[place]	OCP[stop]	IDENT[place]	IDENT[stop]
a. [☞] h pak'ik			*	*
b. [☹] k pak'ik		*!		

Candidate (b), in which no alternation is observed, incorrectly loses due to the fatal violation of OCP[stop], despite the fact that this is the actual output (it is therefore marked by “☹” as a candidate that loses unjustly). From this tableau, we must conclude that OCP[stop] must be lower-ranked than IDENT[place] and IDENT[stop].

Let us examine one further datum: a fricative and a homorganic stop. In this sequence, no phonological alternation is observed, either.

(11) Wrong result: A fricative and a homorganic stop.

/kääs teyäänoh/	OCP[place]	OCP[stop]	IDENT[place]	IDENT[stop]
a. [☞] käähteyäänoh			*	
b. [☹] kääs teyäänoh	*!			

Again, candidate (b), in which no phonological alternation is observed, incorrectly loses due to the fatal violation of OCP[place]. From this tableau, it is concluded that OCP[place] should not outrank IDENT[place].

Thus, the ranking from tableau (9) requires that at least one of the OCPs outrank IDENT[place] and IDENT[stop]. However, OCP[stop] cannot outrank the faithfulness constraints based on tableau (10), and OCP[place] cannot outrank the faithfulness constraint based on tableau (11). We must therefore conclude that there is no valid ranking here to explain all the data above.

2.3 Introduction of a Local Conjunction: OCP[place]&OCP[stop]

Next, we will re-examine the data discussed in the former section by introducing a local conjunction. The local conjunction OCP[place] & OCP[stop] will play a role as a constraint. It will be violated only when *both* OCP[place] and OCP[stop] are violated, not if only one of the members is violated.

Let us examine the conjunction "OCP[place] & OCP [stop]" on the basis of the idea discussed in the first section. The "&" operator in this constraint is in UG, while the choice of OCP [place] and OCP [stop] is specific of Yucatec Maya. The two constraints to be conjoined belong to the same OCP family. Thus, the local conjunction of "OCP[stop] & OCP [place]" is considered to be valid.

The revised ranking which we will propose is as follows:

- (12) OCP[place]&OCP[stop] >> IDENT[place], IDENT[stop]
>> OCP[place], OCP[stop]

With this ranking, let us reexamine the data in tableaux (9), (10), (11), repeated here as (9a), (10a), and (11a):

(9a) A stop and a homorganic stop (revised version of tableau (9)).

/k kool/	OCP[place]&OCP[stop]	IDENT[place]	IDENT[stop]
a. h kool		*	*
b. k kool	*!		

Candidate (b) violates the conjunction because it violates both OCPs. This violation fatally penalizes the candidate. Thus, candidate (a) correctly wins. Candidate (b) also violates the single OCP constraints, but these are ranked lower than IDENT [place] or IDENT[stop] and will not be shown.

Next, let us reexamine the sequence of a stop and a non-homorganic stop:

(10a) A stop and a non-homorganic stop (revised version of tableau (10)).

/k pak'ik/	OCP[place]&OCP[stop]	IDENT[place]	IDENT[stop]
a. h pak'ik		*!	*
b. k pak'ik			

Candidate (b) does not violate the conjunction, since it violates only the lower-ranked OCP[stop] (not shown in the tableau), and wins correctly. Thus, we can obtain the correct output by introducing the conjunction in this tableau in contrast to tableau (10).

Let us now examine another sequence of a fricative and a homorganic stop:

(11a) A fricative and a homorganic stop (revised version of tableau (11)).

/kàas teyàanoh/	OCP[place]&OCP[stop]	IDENT[place]	IDENT[stop]
a. kàahteyàanoh		*!	
b. kàas teyàanoh			

Candidate (b) in this tableau does not violate the conjunction, either, because it violates only OCP[stop]. Thus, candidate (b) is the optimal output. Again, the violation of the single OCP[place] by candidate b is irrelevant (and it is not shown in the tableau), since it is ranked lower than the two IDENT constraints.

We have observed that neither of the two single OCP constraints should be higher-ranked than the faithfulness constraints in the language. Nevertheless, some kind of OCP constraint must account for the phonological alternation in the sequence. Thus, the conjunction OCP[place]&OCP[stop] is necessary to explain the Yucatec data.

The alternations occur so as to satisfy the conjunction, OCP[place]&OCP[stop]. That is why the phonological alternation is observed neither in the sequence of a stop and a non-homorganic stop, nor in the sequence of a fricative and a homorganic stop. In those sequences, only one of the OCP constraints is violated and single OCP constraints can be violated without consequences, since they are lower-ranked.

On the other hand, the conjunction should be satisfied in a sequence of a stop and a homorganic stop (or affricate) at the expense of the violation of the faithfulness constraints, since the conjunction is higher-ranked. Then, the deletion of the [stop] feature, resulting in spirantization, is observed in the sequence. Without the conjunction, we could not account for the phenomena.

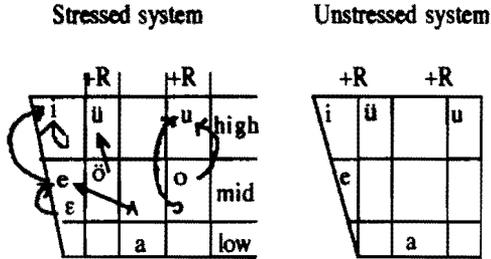
3. Mantuan Vowel Raising in Unstressed Environments

Mantuan vowels present an interesting interaction between universal markedness and faithfulness to underlying forms in unstressed environments. Unstressed syllables lack enough prominence for vowels to surface faithfully to the input. Once a suffix is added to a word, making the stress shift further to the right, away from the original syllable, some features such as ATR and height are changed from the input to the output. In these cases, a more marked vowel that is allowed to surface in stressed positions actually surfaces as a less marked one. The change is driven by the universal markedness of mid vowels (Maddieson 1984, Ladefoged & Maddieson 1990, Beckman 1995 among others) and is perfectly in tune with the "emergence of the unmarked" hypothesis (McCarthy & Prince 1993). The Mantuan vowel neutralization in unstressed environments results in a chain shift where vowels tend to surface as higher vowels, culminating in [i], [u] or [u] where this output is not too unfaithful to the characteristics of the input.

The front vowel chain shift shows that there are two degrees of markedness between mid vowel types: mid vowels are universally marked, but mid lax vowels are more marked than mid tense ones. Thus, a mid lax front vowel never surfaces in Mantuan, but a tense one can. At this point dialectal differences emerge, as in one dialect of Mantuan (Southern Mantuan, "SM") an unstressed mid tense [e] surfaces faithfully. Northern Mantuan ("NM"), however, has a lower tolerance for all mid

vowels, so that an input [e] is too marked to surface faithfully, and surfaces as [i]. This intolerance, however, is masked by the fact that the mid lax vowel surfaces not as [i], but as [e], a move that can be described in terms of opacity (see Kirchner 1996 for a similar case).

Northern Mantuan vowel alternations:



In procedural terms, the phenomenon can be described as “raising” (diagram above): the mid tense vowels raise to high tense ones, maintaining their rounding and back values. Lax vowels can never surface in unstressed syllables, and they therefore acquire an ATR feature, generally maintaining their height distinction.

Mid back vowels, both lax and tense, reduce to [u]. The low vowel remains impervious to any change. Only unrounded front vowels will be taken into account in the present paper (a more complete analysis can be found in Miglio 1996).

3.1 Front Vowel Raising in Unstressed Environments

Mantuan presents the following alternations in unrounded front vowels:

/ε/ as unstressed surfaces as /e/

a)	p_εel	peł'zina	‘skin’~‘cuticle’
	f_εn'estra	fneł'trina	‘window’~‘small window’
	p_εest	peł'tifar	‘pest’~‘mischievous’

/e/, /u/ as unstressed surface as /i/

b)	NM	pél prét pütin	pilín ‘hair’~‘small hair’ prítón ‘priest’~‘big priest’ pütínfn ‘boy’~‘little boy’
b')	SM	pél prét pütin	pelín ‘hair’~‘small hair’ pretón ‘priest’~‘big priest’ pütínfn ‘boy’~‘little boy’

Following Beckman's proposal of positional faithfulness constraints for Shona (1995:60), one can posit a family of constraints on stressed vowels, as separate from a family of general faithfulness constraints. These latter apply in fact to all vowels, and as such, also to unstressed ones. In the following tableaux, it will be assumed that the constraints on stressed vowels are always ranked higher than those on unstressed ones, and will not be considered. The constraints catering for these alternations are modelled after McCarthy and Prince's Correspondence Theory (1995):

- (A) IDENT [height]
Correspondent segments of the Input and Output have identical values for the feature [height] (abbreviated in the tableaux as "ID[hj]" if needed).
- (B) IDENT [ATR]
Correspondent segments of the Input and Output have identical values for the feature [ATR] (abbreviated in the tableaux as "ID[ATR]" if needed).

Since it was observed above that, in this language, the markedness of mid vowels drives some of the alternations, a markedness constraint against mid vowels, *MID, should be posited.

- (C) *MID
All mid vowels are marked.

And finally, since an asymmetrical behaviour between lax and tense mid front vowels can be noticed, to the effect that lax mid vowels never surface in the language, a further, more restricted constraint against mid lax vowels has to be posited, and will have to be placed above the general constraint against mid vowels:

- (D) *MID[-ATR]
All mid lax vowels are marked.

Since there are alternations such as /pel+zína/ surfacing as [pelzína], and /pret+ón/ as [pretón] in SM, and as [prítón] in NM, the analysis will consider first the dialect that only presents the [ɛ] > [e] change, but not [e] > [i], i.e. SM (Quistello dialect). To make data easier to follow, the vowel that presents alternations is underlined in the input and is the only one considered in the outputs shown in the tableaux. They all surface in an unstressed syllable.

3.2 Southern Mantuan (Quistello)

This dialect is characterized by its faithfulness to height values for unstressed /e/: it surfaces as /e/ rather than /i/.

In fact, in tableau (1), a mid[-ATR] input can never surface faithfully (candidate 1 is immediately weeded out), and cannot surface as [i], since it would be violating faithfulness to height (candidate 3). Notice that Ident[height] has to be ranked at least higher than *Mid, but it is unranked with respect to Ident[ATR], this latter is, in turn, unranked with respect to *Mid (a fact that might be obscured in

the tableaux):

(1) Faithfulness to height

pg zina	*Mid[-ATR]	Ident[ATR]	Ident[height]	*Mid
1. ε	*!			*
2. e		*		*
3. i		*	*!	

Candidate 1 is immediately weeded out by the top ranked constraint, and candidate 3 accumulates two violations of faithfulness constraints, which allow candidate 2 to surface.

Tableau (2) shows that an /e/ input surfaces faithfully in SM.

(2) Mid front tense vowel input

pg fn	*Mid[-ATR]	Ident[ATR]	Ident[height]	*Mid
1. ε	*!	*		*
2. e				*
3. i			*!	

The ranking obtained so far for SM is as below:

Southern Mantuan Ranking:

*MID[-ATR] >> IDENT[ATR], IDENT[height] >> *MID

3.3 Local Conjunction in Northern Mantuan (Commessaggio)

Where in SM an input /e/ in a stressed syllable surfaces faithfully in an unstressed one, in NM this surfaces as [i]. This could indicate that a mid vowel is too marked in NM to surface faithfully. Given this line of thought, one would expect NM to have no mid vowels in unstressed positions, but the fact that unstressed [e] can be found shows that the shift driven by markedness is not so clear-cut in this dialect.

The solution to this problem lies in examining the input for a surface [e] in unstressed positions: this input results in fact from a mid lax vowel [ɛ]. Thus, it can be proved that the opacity of this chain shift is to be explained by the fact that a form with a marked vowel is to be preferred, if avoiding the marked vowel means that the surface form is "too different" from its corresponding input. This can be formalized in OT by the introduction of a "last resort" device, which can only apply in such limited situations: NM presents in fact a local conjunction of faithfulness constraints.

The basic ranking that modelled SM front vowels, will not yield the right results in the Northern dialect, cf. (4) and (5):

(4) SM hierarchy for a NM lax input: correct result.

p_ɛ lʒína	*MID[-ATR]	IDENT[ATR]	IDENT[height]	*MID
1. ɛ	*!			*
2. ɛ e		*		*
3. i		*	*!	

Here, given the SM ranking, the front mid lax vowel would be correctly predicted to surface as a mid tense vowel (candidate 2). However, if the hypothesis that this dialect has less tolerance for mid vowels is correct, one is to wonder whether *Mid is too low-ranked to express the higher markedness of mid vowels in NM. In fact, in tableau (5), the problems start to emerge: if [e] is tolerated in unstressed positions, as in the previous tableau, there seems to be no reason why an input [e] should not surface faithfully, as candidate 1 in tableau (5); therefore candidate 1 wins, where candidate 2 should.

(5) Wrong result: SM hierarchy for a NM input.

prɛtón	*MID[-ATR]	IDENT[ATR]	IDENT[height]	*MID
1. ɛ e				*
2. i i			*!	
3. ɛ	*!			*

Thus, as predicted above, the constraint that states the markedness of all mid vowels, *Mid, has to be promoted over the two faithfulness constraints: a crucial difference between Northern and Southern Mantuan. This, however, only gets the right candidate to win in (5), but would lose the winning candidate in (4), modified here as (4a) and (5a):

(4a) Wrong result: Promoting *Mid over faithfulness in NM.

p_ɛ lʒína	*MID[-ATR]	*MID	IDENT[ATR]	IDENT[height]
1. ɛ	*!	*		
2. ɛ e		!*	*	
3. i i			*	*

In tableau (4a) candidate 3 wins, where candidate 2 should.

(5a) Promoting *Mid over faithfulness in NM: this yields the right result in NM.

pre.fm	*MID[-ATR]	*MID	IDENT[ATR]	IDENT[height]
1. ε	*!	*		
2. e		!*		
3. i				*

The right insight should take into account that in NM mid vowels are less tolerated, and that it is better to have *Mid promoted over faithfulness to the input: candidate 1 and candidate 2 are therefore immediately weeded out. Candidate 3 surfaces in spite of the one violation of the faithfulness to height constraint. What should be re-examined, however, is the previous tableau, (4a), where it can be seen that candidate 3 violates both faithfulness constraints and should lose, unlike candidate 3 in (5a) which violates only one and is the right winner.

Therefore, the situation can be summed up as follows: violating two faithfulness constraints in this language is in fact infinitely worse than violating one of them at a time, as well as being worse than surfacing as a marked vowel.

The right way of analyzing Northern Mn is therefore reached by introducing a local conjunction violation (Smolensky 1993, 1995) involving the faithfulness constraints for [height] and [ATR]. "IDENT[ATR]&IDENT[height]" is crucially ranked higher than *MID, as well as than the single IDENT[ATR], and IDENT[height] constraints. These latter are ranked lower than the other constraints in the above tableau: in general this means that in NM a marked input vowel will be likely to surface with different values for ATR and height. As a conjoined constraint, however, IDENT[ATR] & IDENT[height] is higher-ranked than either of them, as well as being ranked higher than *Mid. This will ensure that, in those special situations where the conjoined constraint is violated, a marked mid vowel will be allowed to surface. Local conjunction can be seen at work in the NM revised ranking shown in tableau (6):

(6) Northern Mantuan Local Conjunction: conjunction active.

p e lzína	*MID[-ATR]	IDENT[ATR]&IDENT[hi]	*MID	IDENT[ATR]	IDENT[hi]
1. ε	*!		*		
2. e			*	*	
3. i		!*		*	*

Here candidate 1 loses because of the top-ranked constraint. Candidate 3 violates both faithfulness constraints with respect to the input values: as such this candidate fatally violates the higher ranked local conjunction. This tableau shows why an otherwise marked mid vowel is allowed to surface in this dialect.

In fact, tableau (7) further shows that the conditions under which a marked vowel is allowed to surface in NM are very restricted: candidate 2 violates *Mid fatally and loses in favour of candidate 3, which, given this input, violates one of

the faithfulness constraints, but not the local conjunction.

(7) Northern Mantuan Local Conjunction: conjunction inactive.

pretín	*MID[-ATR]	ID[ATR]&ID[hi]	*MID	IDENT[ATR]	IDENT[hi]
1. ε	*!		*		
2. e			*!		
3. i					*

The ranking for NM should therefore be:

*MID[-ATR], IDENT[ATR]&IDENT[hi] >> *MID >> IDENT[ATR], IDENT[hi]

Thus, for both dialects it is important to rule out mid lax vowels (*MID[-ATR]) before mid tense ones. In NM, on the other hand, all mid vowels are tendentially more marked than in SM, although a difference between lax and tense ones is present there too. In SM mid lax vowels never surface, but a mid tense front vowel [e] is allowed to surface faithfully.

On the other hand, in NM, lax or tense mid vowels never surface, unless a higher local conjunction of faithfulness constraints is violated. The existence of local conjunction in NM obscures the effects of the markedness of mid vowels, by allowing [e] to surface sometimes in unstressed positions. Thus the NM chain shift can be said to be opaque.

Local conjunction should be seen as a "last resort" device, used only in very limited situations. Moreover, it is language-specific: this is shown by the fact that a closely related dialect such as SM bears no trace of it, although it has the single constraints needed to make up the local conjunction of faithfulness constraints.

4. Conclusion

Local conjunction is a device which is made available by UG, through an "&" operator. We propose, however, that the choice of constraints to be conjoined is language-specific, so as to relieve UG from having to encode all conjunctions, as well as all single constraints. The following chart summarizes previous research on local conjunction:

(1) General Chart of Local Conjunctions:

Researcher	Language	Conjoined Constraint	Constraint Family
McCarthy (1996)	Southern Palestinian Arabic	*RTR [HI] & *RTR [FR]	*RTR
Kirchner (1996)	Nzebi	PARSE [LOW] & PARSE [ATR]	PARSE
Miglio (1996)	Mantuan dialect of Italian	IDENT [+ATR] & IDENT [HI]	IDENT
Fukazawa (1996)	Yucatec Maya	OCP[place] & OCP[stop]	OCP

Self Conjunction:

Researcher	Language	Self-Conjoined Constraint	Constraint Family
Alderete (1996)	Wellagga (Oromo)	No Long Vowel & No Long Vowel	No Long Vowel

Constraint Disjunction:

Researcher	Language	Conjoined Constraint	Constraint Family
Hewitt & Crowhurst (1995)	Diyari	Align(Morpheme,L,Foot,L) & Align(Morpheme,R,Foot,R)	Alignment

All researchers in the above chart indicate the need to introduce local conjunction. We have set off the fact that all these conjunctions also belong to the same constraint family⁴. Since no empirical evidence of a local conjunction from two different families (for instance "Max [hi]&*RTR [hi]") has been observed, it is assumed that the restriction on constraint conjoinability to the same constraint family is very strong.

This is a positive result, since, given the architecture of a theory based on strict dominance, it is desirable that a device such as local conjunction be introduced only sparingly, as a "last resort", and in a constrained manner. These characteristics of local conjunction seem to be confirmed by the rarity with which it is encountered, and by the fact that it seems to be constrained by certain conditions, such as locality or phonetic grounding. Based on our independent research (Fukazawa 1996, Miglio 1996), as well as on all previous analyses cited above, we have proposed that another restriction on local conjunction is that constraints belong to the same constraint family.

Notes

1. McCarthy (1996) and Kirchner (1996) use the definition illustrated in this paper, unlike the other researchers mentioned. In the definition of conjunction above, both constraints must be violated in order to violate conjunction. On the other hand, in Hewitt & Crowhurst's definition (1995), local conjunction can be violated whenever *at least one* of the two constraints to be conjoined is violated. Therefore, although part (1) of the definition in this paper also describes their type of conjunction, part (2) differs from theirs.

Alderete (1996) uses "self-conjunction": violating one constraint twice in the same domain is infinitely worse than a single violation of it. Alderete's conjunction is illustrated as follows: A&A >> B >> A. In this sense, Alderete's self-conjunction is closer to the definition in (2). The only difference being that his is a conjunction of the same constraint, and (2) of two different constraints.

2. For further discussion on changing the place feature, see Fukazawa (1996).

3. For [place] as a feature, see Fukazawa (1996).

4. Smolensky (1993) suggests the possibility of a conjunction "NoCoda & *Lab" in order to account for the coda condition proposed by Ito (1986). However, if there are other ways to account for the coda condition, as Ito and Mester (1994) do in their analysis, the need for such a conjunction is open to question.

Bibliography

- Alderete, J. 1996. Self-Conjunction under Strict Adjacency. Ms., University of Massachusetts, Amherst.
- Beckman, J. 1995. "Shona Height Harmony: Markedness and Positional Identity". In University of Massachusetts Occasional Papers: Papers in Optimality Theory. UMass, Amherst, 53-75.
- Fukazawa, H. 1996. An OT Analysis of Yucatec Maya Spirantization. In J. C. Castillo, V. Miglio and J. Musolino (eds.) University of Maryland Working Papers in Linguistics, Vol. 4, Department of Linguistics, University of Maryland at College Park.
- Goldsmith, J. 1976. Autosegmental Phonology. Ph.D. Dissertation. MIT, Cambridge.
- Hewitt, M. and M. Crowhurst. 1995. Conjunctive Constraints and Templates in Optimality Theory. To appear in the Proceedings of NELS 26.
- Kirchner, R. 1996. Synchronic Chain Shifts in Optimality Theory. *LI* 27.2:341-350.
- Ito, J. 1986. Syllable Theory in Prosodic Phonology. Ph.D. Dissertation, University of Massachusetts, Amherst. [New York, Garland Press, 1988]
- Ito, J. and A. Mester. 1994. Reflection on CodaCond and Alignment. In J. Merchant, J. Padgett, and R. Walker (eds.), *Phonology at Santa Cruz III*, 27-46.
- Ladefoged, P. and I. Maddieson. 1990. Vowels of the World's Languages. *Journal of Phonetics* 18, 93-122.
- Leben, W. 1973. Suprasegmental Phonology. Ph.D. Dissertation. MIT, Cambridge.
- Lombardi, L. 1990a. The Nonlinear Organization of the Affricate. *NLLT* 8:375-425.
- Lombardi, L. 1990b. On the Representation of the Affricate. *UMOP* 13, GLSA,

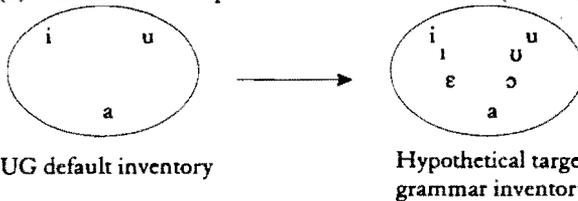
- Amherst, MA.
- Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- McCarthy, J. 1986. OCP Effects: Gemination and Antigemination. *LI* 17:207-63.
- McCarthy, J. 1996. *Process-Specific Constraints in Optimality Theory*. Ms., University of Massachusetts, Amherst.
- McCarthy, J. and A. Prince. 1993. *The Emergence of the Unmarked: Optimality Theory in Prosodic Morphology: Constraint Interaction and Satisfaction*. MS, University of Massachusetts, Amherst and Rutgers University, New Brunswick.
- McCarthy, J. and A. Prince. 1995. *Faithfulness and Reduplicative Identity*. UMOP 18, GLSA, Amherst, MA.
- Mester, A. 1986. *Studies in Tier Structure*. Ph.D. Dissertation, distributed by Graduate Linguistics Student Association, University of Massachusetts, Amherst.
- Miglio, V. 1996. *Mantuan Vowel Shifts*. In J. C. Castillo, V. Miglio and J. Musolino (eds.) *University of Maryland Working Papers in Linguistics*, Vol. 4, Department of Linguistics, University of Maryland at College Park.
- Prince, A. and P. Smolensky. 1993. *Optimality Theory, Constraint Interaction in Generative Grammar*. MS, Rutgers University, New Brunswick and University of Colorado, Boulder. [MIT press, to appear].
- Smolensky, P. (1993). "Harmony, Markedness, and Phonological Activity." Paper presented at Rutgers Optimality Workshop-1, Rutgers University, October 23, 1993.
- Smolensky, P. (1995). "On the Internal Structure of the Constraint Component *Con* of UG." Colloquium presented at UCLA, April 7, 1995.
- Straight, H.S. 1976. *The Acquisition of Maya Phonology: Variation in Yucatec Child Language*. Garland, New York.

Phonological Underspecification and the Subset Principle
 Mark Hale & Charles Reiss
 Concordia University, Montréal

I. The Subset Principle

In this paper we present a radically different view of the Subset Principle (SP) as applied to phonological acquisition from that found in the literature (e.g., Calabrese 1988, 1995, Rice 1996).¹ This traditional view, which we will reject, continues the same basic attitude towards children's speech output as expressed by Jakobson (1941). A standard version of this view is sketched in (1): the initial state of the grammar contains a limited number of vowels, that is, a single vowel or the three 'basic' vowels represented here; acquisition of a larger inventory leads to a superset of this initial inventory.

(1) The Subset Principle in the traditional model (to be rejected)



We consider the relevance of the SP to acquisition to be beyond question, once the assumption is made that children are not sensitive to negative evidence in the course of acquisition. This seems most plausible in phonology where explicit correction of pronunciation errors *vis-à-vis* the target form have long been known to be in vain:

“...these studies show that by the time infants are starting productive use of language they can already discriminate almost all of the phonological contrasts of their native language. While they cannot yet produce adult-like forms, they appear, in many respects, to have adult-like representations, which are reflected, among other things, in their vociferous rejections of adult imitations of their phonologically impoverished productions “ (Faber and Best 1994: 266-7).

In other words, the SP can be viewed as a corollary to the acquisition principle of ‘no negative evidence’. The effect of the SP is to prevent the learner from making overly-broad generalizations which cannot be corrected on the basis of negative evidence alone. We take the essence of the SP to be, therefore, a kind of restrictiveness. In other words, the initial state of the grammar, S_0 , is maximally restric-

tive, and learning consists of relaxing restrictions. Our task, then, is to figure out how these restrictions are formulated (in terms of features, parameters, etc.).

Despite the fact that the SP was first formulated for phonology (Dell 1981) it has been more widely discussed in the syntactic acquisition literature, for example by Berwick (1985) and Wexler and Manzini (1987). Therefore, it may be useful to first review how the SP has been applied to a syntactic problem, as a leadup to our reinterpretation of the principle. Given our concerns, the discussion of syntactic phenomena will be extremely informal.

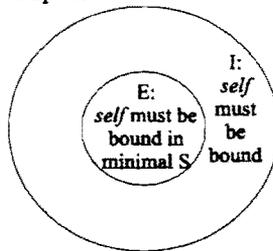
In a comparative study of acquisition of anaphora, Hyams and Sigurjónsdóttir (1990) compare the binding conditions on Icelandic *sig* and English *him/herself*. In simple terms, we can characterize the anaphors in the two languages as follows: Icelandic anaphors need to be bound; English anaphors need to be bound in the minimal clause (S). So, English is more restrictive, it imposes more conditions on anaphors than Icelandic does. The difference is represented by the schematic sentences in (2). In English, the anaphor can only be coreferential with the NP in the same clause, whereas in Icelandic, the anaphor can be coreferential with an antecedent in a higher clause.

(2) Anaphors in English and Icelandic

- English: John_i asked Bill_j to shave self*_{i/j}
- Icelandic: John_i asked Bill_j to shave self_{i/j}

We can represent the greater restrictiveness of English as in (3) and conclude that English corresponds to the initial state (in this respect).

(3) The subset principle for anaphors



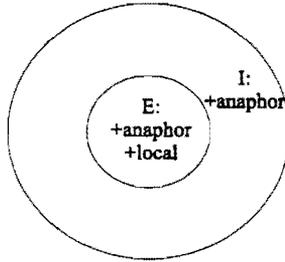
We can also represent the relationship of the two languages as an implicational relationship, as in (4).

(4) The SP as an implicational hierarchy

- a) Anaphor must be bound in the minimal S \Rightarrow Anaphor must be bound.
- b) Anaphor must be bound \Rightarrow Anaphor must be bound in minimal S.

If we try to represent the distinction between the languages in terms of lexical features, instead of in terms of parameter settings, as has been done traditionally, we might propose the model in (5), where English anaphors are marked as [+anaphor, +local] whereas Icelandic anaphors are marked only as [+anaphor].

(5) Features for anaphors



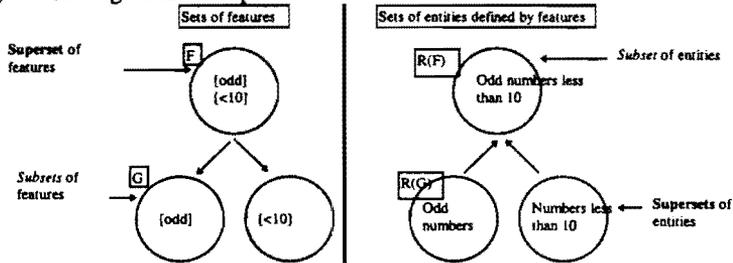
This brings us to a point, which though obvious, is crucial to our argument. Despite the simplicity of the argument, it is precisely the failure to grasp this principle which has led to the misinterpretations of the SP which we will present below.

The relevant point can be formulaically stated as: “fewer features equals more things”. That is, the size of a class varies inversely with the number of features used to define the class. This is stated more formally in (6).

(6) Let F and G be sets of features such that $R(F)$ is the set of entities defined by F and $R(G)$ is the set of entities defined by G . If G is a subset of F , then $R(F)$ is a subset of $R(G)$. That is $F \supset G \Leftrightarrow R(G) \supset R(F)$.

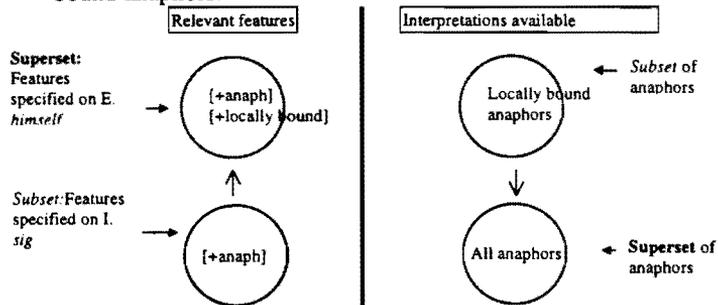
At the risk of appearing pedantic, we now present a non-linguistic example of this principle. The properties of being ‘odd’ and being ‘less than 10’ can be used to characterize, positively or negatively, the set of whole positive numbers. As shown on the left-hand side of (7), the set of properties, or features, containing both ‘odd’ and ‘< 10’ contains the two sets which contain only one of these features. On the right-hand side, however, we see that the containment relation goes in the other direction: the set of numbers which are both odd and less than 10 is contained within the set of odd numbers and within the set of numbers less than 10.

(7) A non-linguistic example



We can now return to our linguistic example and see that the same inverse relation holds. On the left-hand side of (8) we see a superset of features containing a subset of features, but on the right hand side we see that the interpretations associated with anaphors are in the inverse relationship.

(8) Linguistic example: the class of anaphors contains the class of locally bound anaphors:



We now turn to discussion of some phonological cases from the literature, both concerning the acquisition of metrical phonology.

Archibald 1995 proposes that the SP is not relevant to the acquisition of English extrametricality.

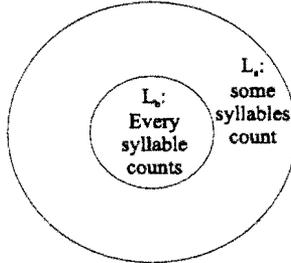
(9) Archibald 1995 on the Subset Principle:

the two possible settings of the extrametricality parameter [+/- that is—mrh & cr] do not yield languages that are in a subset relation to each other. If English has no extrametricality ([-extrametrical]) we would always find penultimate stress (assuming left-headed feet). If it had obligatory extrametricality ([+extrametrical]) we would always find antepenultimate stress. The situation is a little more complex, as we have extrametricality in English that is sensitive to the grammatical category (i.e. it works differently on nouns and

verbs, for instance). Thus, the Subset Principle makes no clear predictions.(86)

We propose an alternate view, namely that the initial hypothesis made by the child, the initial setting of an 'extrametricality parameter', is that **every syllable must be considered for the computation of stress**. The initial grammar, then, is **more restrictive** than any subsequent grammar which may, as in the case of English, allow certain syllables to be left out of stress computation (either based on independent computations such as foot construction or through lexical marking). This initial language, or UG, is thus true to the essence of the SP. In (10) then, L_b corresponds to the initial state, S_0 . Again, we can state the relationship between S_0 and later states as an implicational one.

(10) The SP and extrametricality



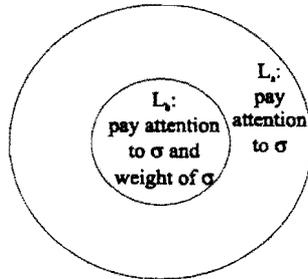
- a) Every syllable counts \Rightarrow Some syllables count.
 b) Some syllables count \nRightarrow Every syllable counts.

Our next example comes from Dresher 1996 who assumes that sensitivity to syllable weight must be acquired. Dresher's claims are sketched in (11), where σ_H is a heavy syllable and σ_L is a light syllable.

(11) Dresher (1996) claims that sensitivity to syllable weight must be acquired, so a language which does not 'care about weight' is the initial state.

In our view this is backwards, since caring about weight is a restriction on foot construction. Children must be innately sensitive to syllable weight. If children did not have this sensitivity, then positive evidence would never provide it to them, and weight sensitive systems would be unlearnable. In learning a language in which syllable weight does not matter, the child **learns to ignore** a distinction which must be possible in UG. We assume that the subset relation must be the contrary of what Dresher proposes. This means that L_b actually represents the initial state of UG: $L_a: \{\sigma\} \supset L_b: \{\sigma_L, \sigma_H\}$, so L_b is the initial state, as shown in (12), along with another implicational statement.

(12) Our proposal: $L_a:\{\sigma\} \supset L_b:\{\sigma_L, \sigma_H\}$. Contrast with Dresher (1995): $L_a:\{\sigma\} \subset L_b:\{\sigma_L, \sigma_H\}$



- a) Paying attention to syllables and their weight \Rightarrow paying attention to syllables.
 b) Paying attention to syllables $\not\Rightarrow$ paying attention to syllables and their weight.

II. SP and segment 'inventories'

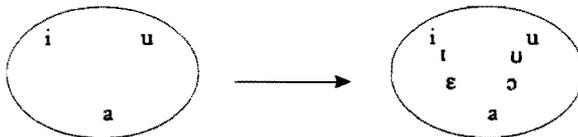
We now turn to the main topic of this paper, the acquisition of phonological inventories. One of the most explicit versions of the traditional model is presented by Rice (1996), sketched in (13).

(13) The Subset Principle in the traditional model (rejected)

Central hypotheses (Rice 1996):

Minimality: Initially the child has minimal structure (not all features are available.)

Monotonicity: Inventories are built up in a monotonic [stepwise—mrh&cr] fashion as contrasts are added.



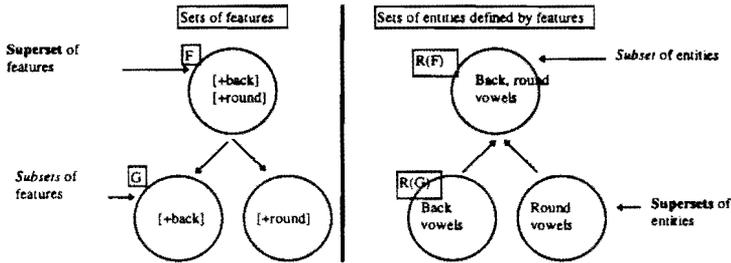
UG default inventory

Hypothetical target
grammar inventory

We can object to this hypothesis on two grounds. First, there is no reason to expect 'segments' to play a role in the learning path, since features are the primitives of phonological theory. Second, the inverse relationship of features and natural classes discussed above leads to an alternative interpretation.

(14) Objections

- Languages (grammars) do not have /i/, /a/, /k/, etc. The “atoms” of phonology are features (cf. the *wug* test, etc).
- Fewer features = more entities. The class of back vowels contains the class of back, rounded vowels:



With these objections in mind, consider the ‘phonological space’ associated with vowels in two languages, one with a rich inventory and one with a restricted inventory, shown in (15). The direction of the subset/superset relationship is not so clear when faced with two ways of looking at the problem: i) numbers of ‘segments’ and ii) phonological space.

(15) Phonological space assigned to high front vowels in two vowel systems: which is the subset?



The arguments we have offered to this point favor choosing the language with more restrictive, i.e. richer, representations and narrower phonological space associated with individual vowels as the initial state. In order to provide learnability arguments to support this proposal and justify rejecting the traditional theory, we must answer the two questions in (16). Below we provide arguments using hypothetical languages to justify the answers we provide.

(16) The tests:

- Can the traditional view lead to a growing inventory? NO.
- Can the proposed view lead to a shrinking inventory? YES.

In order to answer (16a) consider the acquisition of /dip/ vs. /dɪp/ in a hypothetical language which maintains the [i] / [ɪ] contrast on the surface. In the traditional system, the contrast is unlearnable, the two words will be acquired as homophones. Without access to a difference in representation, the difference between the two vowels cannot be evaluated. The so-called 'positive evidence' often invoked to allow inventory expansion is not sufficient if that evidence cannot be assigned a representation. That is, the contrast cannot be processed linguistically if the child doesn't have the contrast. This is a fundamental assumption of linguistic theory. It is equivalent to saying that a language that uses a non-human feature is unlearnable, which is tautologically true given the standard definition of UG. If a child did not have access to a feature provided by UG, then the child could not store this distinction for future use; each lexical entry would have to be relearned at each stage since each lexical entry could potentially contain the newly 'acquired' feature. This is contraindicated by the acquisition evidence. On the other hand, if the distinction is available at S_0 then acquisition of contrastive lexical items is trivial. Our claim is consistent with the evidence from comprehension and perception studies (Streeter 1976, Faber and Best 1994) and thus should represent the null hypothesis.

We now turn to (16b), loss of a 'wrong' contrast, i.e. /dip/ and /dɪp/ collapse to /dɪp/ in some language with a three vowel system. The traditional view will never face a problem here, since the grammar never contains more contrasts than the target language. The challenge to the theory proposed can be stated thus: How does a grammar which has more potential vowels than the target grammar end up losing irrelevant contrasts? Two cases must be distinguished.

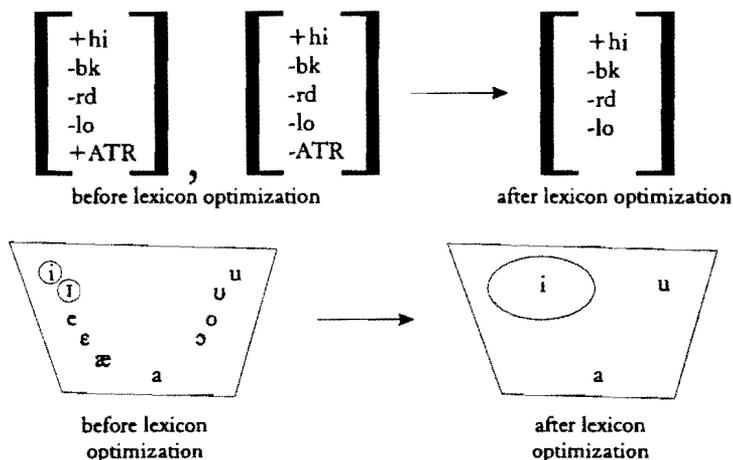
(17) Two distinct cases for the 'collapse' of contrast

- a. Unobserved contrasts: If the target language does not present forms such as [dip], then there will never be any reason to remove [+ATR] from the representation of /dip/. Access to the universal feature set allows the *potential* for any contrast, not its realization. This is clearly relevant to underspecification theory in that we assume, for instance, that /i/ will never lose its [+ATR] specification without grammar internal motivation (see below).
- b. Phonetic underspecification: Imagine the child hears [dip] and stores it as such. Since this child has access to all the features and since its learning conforms to the SP, it assumes that representations must be maximally restrictive (specified). This word cannot be stored with just a [-back, +hi] vowel; it must be stored as a [-back, +hi, -round, +ATR] vowel. Given the variability of articulation in a three vowel system, this child will hear phonetic [dɪp] and mistakenly posit a new lexical item, ending up with a pair of

synonyms, /dɪp/ and /dɪp/. A process of lexicon optimization, responsible for collapsing synonyms will determine that the feature is not relevant to the phonology of the language and thus can be excluded from representation. This proposal is very similar to the uniqueness principle of morphology (Wexler and Culicover 1980). (This algorithm does not affect all redundant features, merely those which behave as though they are *phonetically underspecified*.)

Note, at this point that the confusion is largely notational. In losing the /i/-/ɪ/ contrast the grammar moves from containing two vowels [-back, +hi, -round, +ATR] and [-back, +hi, -round, -ATR], which we happen to denote as /i/ and /ɪ/, to one [-back, +hi, -round] which we somewhat arbitrarily denote as /i/.

(18)



Clearly, this account needs to be further developed by an explicit model of lexicon optimization. Crucial questions include the nature of optimization: is it a global process which affects the whole lexicon at once, or does it proceed on a morpheme-by-morpheme basis?

III. Summary

We can summarize the argument to this point with two subparts which lead to the same conclusion in the following way:

(19) Summary of arguments

A. Subset Principle Argument

- a. The Subset Principle reflects restrictiveness in the initial state
- b. Greater restrictiveness is encoded through fuller specification
- c. ∴ All features must be available for representations at S_0 .

B. Learnability Argument

- a. Linguistic representations contain features
- b. If a feature F is unavailable at stage L_j , then positive evidence of F cannot be evaluated by the learner since the learner cannot evaluate representations with respect to F
- c. ∴ All features must be available for representations at S_0 .

We turn now to a consideration of how these results bear on issues in phonological theory in general, especially the status of underspecification.

IV. Inkelas, et al. approach to underspecification

The approach to underspecification advocated in this paper is consistent with recent work by Sharon Inkelas and her collaborators (Inkelas 1994, Inkelas, Orhun and Zoll 1996) in the adoption of the view that underspecification is derived from the data, not imposed by principle. This position, exemplified by (20) leads Inkelas to reject philosophically based approaches to underspecification such as those represented by the (partially overlapping) categories in (21).

(20) "underlying representation is determined solely by optimization with respect to the grammar, not by imposing any type of constraints directly on underlying representation...[this] results in the use of underspecification only when there are alternant surface forms..." (Inkelas 1994:1).

(21) 'Grammar-blind' approaches to underspecification rejected by Inkelas 1994, q.v. for references

- Markedness** (universal, language-specific, or contextual); unmarked material is underspecified
- Redundancy**; redundant feature values (determined on the basis of the segment inventory) are underspecified
- Predictability**: predictable material is underspecified

According to Inkelas, "[t]he only motivation for underspecification is to capture alternations in the optimal way" (1994: 2). A very convincing case where underspecification can be used in an illuminating fashion comes from the distribution of Turkish voiced and voiceless stops. As the data in (22) show, the two Turkish surface stops [t] and [d] show up in three different patterns.

(22) Turkish voicing alternations

- a. Alternating: [0voice] (unmarked for [voice])
kanat 'wing' kanatlar 'wing-plural' kanadım 'wing-1sg.poss'
- b. Non-alternating voiceless: [-voice]
sanat 'art' sanatlar 'art-plural' sanatım 'art-1sg.poss'
- c. Non-alternating voiced: [+voice]
etüd 'etude' etüdler 'etude-plural' etüdüm 'etude-1sg.poss'

In (a), the stop is voiceless in coda position and voiced in onset position. Inkelas proposes that this stop is underlyingly unspecified for [voice] and it receives its specification by the equivalent of structure-filling rules. The (b) forms are stable in always showing voiceless stops and thus are underlyingly specified as voiceless. The (c) forms are consistently voiced, and thus are underlyingly marked as voiced. Inkelas, Orgun and Zoll 1996 argue convincingly that this is the best account of this type of data, rejecting, for example, the use of lexical exception features.

While we accept fully the spirit of this argument, we find it necessary to expand the range of underspecification in an additional way. This is the case of phonetic underspecification in (17b), whereby the grammar just 'doesn't care' about the setting of a certain feature. This was discussed for a hypothetical case above, where the target space for the high, front vowel in a three vowel language included both [+/-ATR] regions. The existence of such articulatory freedom in vowel articulation is an empirical issue supported by some phonetic studies of relevant languages (Manuel 1987, but see Maddieson and Wright 1995 for a potential case of a three vowel system with highly restrictive target spaces). If it turns out, however, that such variation is in fact contextually determined, then these cases of apparent 'phonetic underspecification' will reduce to conditioned allophony.

As a concrete example of how this approach differs from standard views on underspecification consider the acquisition of the lexical representation of the /n/ in English /ten/ 'ten'. Since the child starts out with full specification and removes specification only when 'forced' to by the grammar, representations may be grossly 'overspecified' (even in adult grammar). There is no motivation to leave this segment unspecified for [voice], since no alternations exist which bear on the issue, and the ambient language has voiced nasals exclusively. The theory explicitly rejects the grammar-blind principles in (21), not least on account of the severe empirical difficulties they face, as discussed by Inkelas and her references. The dental/nondental specification, on the other hand, must be dealt with in order to allow the allomorphs in [ten] 'ten' and [teŋθ] 'tenth' to be derived from a single UR. We leave several crucial issues for further research. One concerns the process of lexicon optimization which accomplishes the collapse of the allomorphs in 'ten' and 'tenth'. Does this proceed on a lexeme-by-lexeme basis, or is

the whole lexicon scanned *en masse*? Second, what is the result of this collapsing in a given case? Does it result in specification of one value, say non-dental, and a feature changing rule (or OT equivalent), or does the dental/nondental feature specification get erased from every /n/ and get reassigned by feature-filling rules. Without going into a full discussion, we prefer to reserve feature filling operations for cases such as the Turkish voicing alternations discussed by Inkelas, that is, cases where the grammar requires them. A third question, related to the others, concerns the scope of 'optimization': does it affect only those representations which participate in alternations, or does the lexicon get 'optimized' according to the alternations that exist anywhere in the language? In other words, what happens to the dental nasal of 'plinth', a non alternating morpheme?²

Yip (1996) reaches basically the same conclusion as Inkelas (1994), both of which are compatible with the theory of rich initial representations espoused here:

(23) "This paper finds inconclusive evidence for abstract underlying representations, and concludes that the balance of the evidence suggests that learners acquire something rather close to what they hear, unless information from alternations or paradigms forces them to do otherwise" (Yip 1996).

Both Yip and Inkelas argue that their conclusions follow from the nature of Optimality Theory. In our view, a principled account of acquisition and underspecification can be accomplished in a variety of frameworks.

V. Conclusions and Implications

We conclude with the following list of observations:

- The Subset Principle is a valid principle of acquisition, but formulation in a given case may be contrary to what is often assumed.
- The Subset Principle must be stated with reference to features—more features define smaller sets of representations, more feature specification means more restrictive representations.
- The standard view of phonological acquisition is incompatible with empirical evidence such as perception and comprehension studies (Hale and Reiss 1996a, 1997).
- Smolensky (1996a) agrees with us (and Yip 1996) in assuming that children initially have access to full specification. Therefore 'richness of the base' (e.g. Prince and Smolensky 1993,191; Smolensky 1996b) is irrelevant to acquisition (at least for non-alternating morphemes) and perhaps to human grammar in general. It is merely a computational curiosity of OT grammars. Not surprisingly, the inventory acquired reflects the ambient language. Acquiring the URs and surface forms of the ambient language drives constraint reranking

during acquisition. This can be opposed to the view that the inventory is derived from the constraint ranking (Smolensky 1996b). See Hale and Reiss (1996ab).

- Given that representations must start out fully specified, children tell us that the so-called 'emergence of the unmarked' (McCarthy and Prince 1993) represents a learned, rather than UG default, phenomenon. Children's grammars are thus generally faithful to the adult output which they store. Apparent unfaithfulness must be attributed to their performance systems, not to their grammars. It follows, therefore, that OT Faithfulness constraints must initially be ranked high in children's grammars, *contra* Smolensky (1996a). See Hale and Reiss (1996ab, 1997) and Scobbie et al (1997).
- The model presented in this paper is consistent with current conservative views of underspecification.
- Given current assumptions concerning an invariant syntactic component with cross-linguistic variation effected through lexical feature differences, one might expect that the Subset Principle in syntax would work in the same way as proposed for phonology.

¹ See Reiss (1995) for application of the subset principle to an apparent implicational universal.

² We tend to agree with a comment by Alan Prince posted to the OT discussion list (Nov.21, 1996): "why, in a grammar G such that G(a)=G(b) for potential input elements /a/,/b/, [is] a nonalternating observed element [a] ... not (sometimes, always, freely) lexically /b/? The correct answer is surely 'why bother?' — i.e. to set up /b/ for[a] when /a/ will do. This common-sense approach to choosing URs for non-alternating morphemes is not, however, standard practice in the OT literature. Note, in addition, that Prince's view is tantamount to admitting the irrelevance of 'richness of the base' in cases of non-alternating morphemes (see Hale and Reiss 1996b).

References

- Calabrese, Andrea. 1995. A Constraint-Based Theory of Phonological Markedness and Simplification Procedures. *Linguistic Inquiry* 26: 373-463.
- Calabrese, Andrea. 1988. *Towards a theory of phonological alphabets*. Ph.D. dissertation, MIT.
- Dell, François. 1981. On the learnability of optional phonological rules. *Linguistic Inquiry* 12: 31-37.
- Dresher, Elan. 1996. Charting the learning path: Cues to Parameter Setting. Talk presented at *MOT*, McGill University. Ms. University of Toronto.

- Faber, Alice and Catherine T. Best. 1994. The perceptual infrastructure of early phonological development. In *Reality of Linguistic Rules. Studies in language companion series*; v. 26. Susan D. Lima, Roberta L. Corrigan, Gregory K. Iverson, eds. J. Benjamins, Amsterdam.
- Hale, M. and C. Reiss. 1996a. How to Parse (and How Not to) in OT Phonology. Paper presented at NELS 27, McGill University. Ms. Concordia University; to appear in *Proceedings of NELS 27*.
- Hale, M. and C. Reiss. 1996b. Formal and Empirical Arguments Concerning Phonological Acquisition. Ms. Concordia University; also available on ROA.
- Hale, M. and C. Reiss. 1997. Evidence in Phonological Acquisition: Implications for the Initial Ranking of Faithfulness Constraints. *Proceedings of CLRF 28*. CSLI Publications, Stanford CA.
- Hyams, N. and S. Sigurjónsdóttir. 1990. The development of 'long distance anaphora': a cross-linguistic study with special reference to Icelandic. *Language Acquisition* 1, 57-93.
- Inkelas, Sharon. 1994. The Consequences of Optimization for Underspecification. Ms. UC Berkeley. ROA.
- Manuel, Sharon. 1987. *Acoustic and Perceptual Consequences of Vowel-to-Vowel Coarticulation in three Bantu Languages*. PhD. dissertation, Yale University, New Haven.
- McCarthy, J. 1994. Remarks on Phonological Opacity in OT. Ms. UMass, Amherst. To appear in *Proceedings of the Second Colloquium on Afro-Asiatic Linguistics*. J. Lecarme, J. Lowenstamm and U. Schlonsky, eds.
- McCarthy, John and Alan Prince. 1993. *Prosodic Morphology I: constraint interaction and satisfaction*. Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ. To appear as Linguistic Inquiry Monograph, MIT Press, Cambridge, MA.
- Prince, A. and P. Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical Report Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, N.J.
- Reiss, Charles. 1995. Deriving an implicational universal in two theories of phonology. Ms., ROA-102.
- Rice, Keren. 1996. Phonological Variability in Language Acquisition: A Representational Account. Ms. University of Toronto.
- Scobbie, J., W. Hardcastle, P. Fletcher and Fiona Gibbon. (to appear). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe and J. Pierrehumbert (eds.). *Papers in Laboratory Phonology V*. Cambridge University Press.
- Smolensky, P. 1996a. On the Comprehension/Production Dilemma in Child Language. *Linguistic Inquiry* 27, 720-731.

- Smolensky, Paul. 1996b. The initial state and 'richness of the base' in Optimality Theory. Technical Report JHU-CogSci-96-4, Department of Cognitive Science, Johns Hopkins University.
- Streeter, L. A. 1976. Kikuyu labial and apical stop discrimination. *Journal of phonetics* 4:43-9.
- Wexler, Kenneth and Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.
- Wexler, Kenneth and M. Rita Manzini. 1987. Parameters and Learnability in Binding Theory. Parameter Setting, ed. by Thomas Roeper and Edwin Williams, 41-76. Reidel, Dordrecht.
- Yip, Moira. 1996. Lexicon Optimization in Languages without Alternations. Ms. University of California, Irvine. ROA.

The role of comprehension, reinterpretation and the Uniformity Condition
in historical change: the case of the development of *Cl* clusters from Latin
to Hispano-Romance

D. Eric Holt
Georgetown University

0. Introduction.* As shown in (1), clusters of voiceless consonant and /l/
/k, p, f + V/ undergo a series of changes from Latin to Hispano-Romance:

0.1 Data:

(1)		<i>Latin</i>	<i>Spanish</i>	<i>Galician/Portuguese</i>		
(a)	<i>initial</i>	CL	CLAVE	<i>llave</i> (ʎ)	<i>chave</i> (tʃ)	'key'
			CLAMARE	<i>llamar</i>	<i>chamar</i>	'to call'
		PL	PLUVIA	<i>lluvia</i>	<i>chuva</i>	'rain'
			PLANCTU	<i>llanto</i>		'weeping'
			FLORARE	<i>llorar</i>	<i>chorar</i>	'to weep'
			PLAGA	<i>llaga</i>	<i>chaga</i>	'wound'
			PLICARE	<i>llegar</i>	<i>chegar</i>	'to arrive'
			PLENU	<i>lleno</i>	<i>cheio</i>	'full'
		FL	FLAMMA	<i>llama</i>	<i>chama</i>	'flame'
			FLACCIDU	<i>llacio</i> (later <i>lacio</i>)		'lank'
(b)	<i>medial</i> ¹	CL	* <i>mancla</i> ²	<i>mancha</i> (tʃ)	<i>mancha</i> (tʃ)	'stain'
			CONCHULA	<i>concha</i>	<i>concha</i>	'shell'
			TRUNCULU	<i>troncho</i>		'stalk'
			HINNIT-*ULARE	<i>reninchar</i>	<i>r(el)inchar</i>	'to whinny'
		PL	IMPLARE	(<i>h</i>) <i>enchar</i>	<i>encher</i>	'to fill'
		FL	INFLARE	(<i>h</i>) <i>inchar</i>	<i>inchar</i>	'to inflate'
			* <i>masclo</i>	<i>macho</i>	<i>macho</i>	'male, macho'
			CICERCULA	<i>cizercha</i>		'blue vetch'
			SARCULARE	<i>sachar</i>	<i>sachar</i>	'to weed'
			AFFLARE	<i>hallar</i>	<i>achar</i>	'find, think'

(For exceptions, see Malkiel 1963-4)

The data can be summarized as follows: In both Spanish and Galician/Portuguese in medial position the result is /tʃ/, while in initial position the two languages differ, with Galician/Portuguese showing /tʃ/ (later /ʃ/), but Spanish showing a different outcome, /ʎ/. Previous authors' proposed derivations are in (2).

0.2 Previous accounts. Other researchers have addressed these changes:

- (2) Williams (1938): $Cl > Cj > tʃ$ (only Galician/Portuguese treated)
 Bourciez (1967): $Cl > ll > ʎ > tʃ$
 Lloyd (1987): $Cl > Cʎ > ʎ (> tʃ$ medially; later generalized to initial position in Galician/Portuguese)

The first two authors fail to consider that Upper Aragonese shows /Cʎ/ (where 'C' represents /k, p, f/), and fail to treat Galician/Portuguese as having once shared a stage with Spanish. Lloyd, however, recognizes the importance of the Modern Upper Aragonese data, shown in (3), which allows him to develop a more unified approach:

(3)	Latin	Upper Aragonese ³
CL	CLAVE	<i>cllau</i> [kʎ] 'key'
PL	PLOVERE	<i>pllover</i> [pʎ] 'to rain'
FL	FLAMMA	<i>fllama</i> [fʎ] 'flame'

Nonetheless, all previous researchers assume some sort of 'magic leap' from /*Cʎ/, /*ʎ/ or /*Cj/ to /tʃ/. That is, it is assumed that a voiceless consonant + front semivowel (in the case of Williams), a voiceless consonant + /*ʎ/ (in the case of Lloyd), or just the /*ʎ/ (in the case of Bourciez) develops directly to /tʃ/. However, these are very different sounds, and none of these authors proffers an analysis as to *how* or *why* the situation and change should be as they are. Phonetically, many of these proposed changes are just plain hard to justify given that they assume some kind of articulatory or acoustic gap for which they do not account.

0.3 Principal issues of this paper. The present account provides a unified approach to the various Hispano-Romance dialects, and provides an explanation for the 'magic leap' previously stipulated. In addition, the present account also raises a number of theoretical issues, some of which have not been addressed in Optimality Theory:

- (4) (a) Phonetics → phonology → lexicon (then repeat the cycle)
 (b) The Uniformity Condition played a role in this varied development (in OT via conjunction of constraints and ranking of conjoined constraints)
 (c) The role of the listener in historical change (cf. Ohala, Janson, Jonasson, etc.): Perception and comprehension lead to reinterpretation (here via acoustic equivalence, emergence of the unmarked and lexicon optimization)
 (d) Certain similarity of historical change to child language acquisition and learning algorithms

1. Analysis. Here I wish to maintain the unity of Spanish and Galician-Portuguese, and follow Lloyd in assuming that Upper Aragonese shows an intermediate stage in the development from Lat. *Cl* to OSp., Gal./Ptg. *ch*. I further motivate this change for both Spanish and Galician/Portuguese in medial position, and for Galician/Portuguese in initial position, rather than having [tʃ] be generalized from medial to initial position.

I now offer my analysis of the series of changes that occurred.

The first stage is the assimilation of /l/ to /k/, yielding [*kʎ]. The articulation of /l/ is drawn toward the velar region where /k/ is pronounced. The data from Rumanian (where only the /kʎ/ clusters palatalized, leaving /pl, fl/ as is; see Tuttle 1975, Lloyd 1987, others) are generally taken as supporting the assumption that this is the first step.

- (5) First proposed historical stage: Assimilation. /kʎ/ > [*kʎ]
 (Hispano-Romance, medial position; later also initial position in pre-Old Spanish)

/kʎ/ > /*kʎ/	ASSIMILATE[PA]	IDENT[PA]
kl	*!	
<input checked="" type="checkbox"/> kʎ		*

Hispano-Romance forms:

*MACULA > [*maŋkʎa] > [*maŋkʎa]
 AURICULA > [*orekʎa] > [*orekʎa]

Later, also in pre-Old Spanish:

CLAMARE > [klamar] > [*kʎamar]
 CLAVE > [klave] > [*kʎave]

This begins as a phonetic process, but is then phonologized and lexicalized by the listener.⁴

However, since not only CL but also PL and FL developed to /tʃ/ or /kʎ/, the next stage in this development is the extension of /kʎ/ to /pl, fl/. For these clusters the initial consonant is produced with the lips, not the hard

palate, and therefore there is no phonetic factor that would cause /l/ to become palatal [ʎ]:

- (6) Second proposed historical stage: 'Allophonic unification'.
/pl, fl/ > [*pʎ, *fʎ] by influence of /*kʎ/ (Tuttle 1975:407-8)

/*kʎ/ was the most frequent *Cl* cluster, and as such it could have served as a robust model for analogical change: [*ʎ] is thus extended to /*pʎ, *fʎ/, as in Modern Upper Aragonese *pllover*, *flama*.

The predominant source of /*kʎ/ was by reduction of the diminutive suffix -ICULUS > -CLO, OCULUS > [*okʎo]. Additional examples are given below (I show the complete historical derivation for only the first example):

- (7) /*kʎ/ as model for 'allophonic unification' of /pl/, /fl/ to /*pʎ/, /*fʎ/:

AURICULA (for AURIS) > [*orek'la] > [*orekʎa]	'ear'
OVICULA (for OVIS) > [*ovekʎa]	'sheep'
APICULA (for APIS) > [*abekʎa]	'bee'
CLAVICULA (from CLAVE) > [*k(l)avekʎa]	'peg, pin'
OCULUS > [*okʎo]	'eye'
SPECULUM > SPECTUM > [*espekʎo]	'mirror'
VETULUS > VECLUS > [*vekʎo]	'old'
LENTICULA > [*lentekʎa]	'lentil'
VERMICULU 'little worm' > [*bermekʎo]	'red'
COAGULU > [*koagʎo]	'curds'
REGULA 'metal bar' > [*r:egʎa]	'plowshare'
TEGULA > [*tegʎa]	'roof tile'

I tentatively suggest that 'allophonic unification' may be considered to aid in the economy of lexical representations, and that this kind of sequential constraint is a kind of lexicon optimization. Due to limitations of space, I will have to leave it at that for the present discussion. (I explore this further in Holt in preparation.)

This assimilation applied only word-internally in Hispano-Romance at first, but its application spread to initial position, and did so more quickly in Old Spanish than in Old Portuguese. This is supported by the fact that there is much more variability of outcome in initial position, particularly in Portuguese. (See Wireback 1996 for discussion of the factors involved in

the spread of this sound change. For the 'conservatism' of Galician-Portuguese, see Lloyd 1987, Repetti and Tuttle 1987, Holt in preparation, others.)

However, the articulation of this cluster is quite complex, and it is subsequently reduced. An appropriately-modified version of the following constraint is active:

- (8) *COMPLEX:
No more than one consonant or vowel may associate to any syllable position node.
(Prince and Smolensky 1993:87, Hargus 1995)

The interaction of this constraint with MAX determines the simplification of these clusters. This is the third stage in the historical development treated here:

- (9) Third proposed historical stage: Simplification. /^{*}CK/ > /K/
(Hispano-Romance, most positions; that is, all positions where there were CΛ clusters)

/ [*] CK/ > /K/	*COMPLEX (ONSET)	MAX (SONORANT)	MAX (OBSTRUENT)
CΛ	*!		
C∅		*!	
<input checked="" type="checkbox"/> ∅Λ			*

This occurred medially for both pre-Old Spanish and Galician/Portuguese, as well as for the initial /^{*}CK/ clusters of pre-Old Spanish:

- (10) Hispano-Romance³ Old Spanish
 OCULO > [^{*}okΛo] > [oΛo] CLAMARE > [^{*}kΛamar] > *llamar*
 COAGULU > [^{*}koagΛo] > [koΛo] FLUVIA > [^{*}pΛuvja] > *lluvia*
 (MPtg. *olho, coalho*) FLAMMA > [^{*}fΛama] > *llama*

The loss of the first rather than the second consonant is determined by the ranking of MAX(SONORANT) >> MAX(OBSTRUENT). This ranking is consistent not only with the data described here but also with the general pattern of simplification observed from Latin to Hispano-Romance; another instance of this simplification via loss of the initial obstruent is GL- > /-, BL- > /-:⁶

(11) Simplification of /bl/ and /gl/ to /l-/ in Hispano-Romance:

BLASPHEMARE	>	<i>lastimar</i>	'to damage'
BLATTA	>	OSp. <i>lad-illa</i>	'crab louse'
GLATTIRE 'to bark'	>	<i>latir</i>	'to beat'
GLANDINE 'acorn'	>	OSp. <i>landre</i>	'tumor'
GLIRE	>	OSp. <i>lir</i> (MSp. <i>lirón</i>)	'dormouse'
GLOBELLU	>	OSp. <i>loviello</i> (MSp. <i>ovillo</i>)	'ball [of yarn]'
FABULARE	>	Ptg. <i>falar</i>	'to speak'

To recapitulate the discussion thus far: What begins in Late Latin as assimilatory palatalization of /kl/ to /*kʎ/ is extended by analogy to the medial clusters /pl, fl/ to /*pʎ, *fʎ/ (and to initial position in pre-Old Spanish), and these articulatorily complex clusters are simplified from /*Cʎ/ to /kʎ/.

Here is where we reenter the written record. Written documentation goes from Lat. *-Cl-* to OSp. *-ll-*, OPtg. *-lh-* (= [ʎ]), and from Lat. #*Cl-* > OSp. *ll-*. Also at this historical stage, *ch* ([tʃ]) now appears in medial position in both Old Spanish and Old Portuguese.

To the best of my knowledge, no previous analysis has adequately, if at all, explained why medial position should have developed differently than in initial position. This difference has been observed, but not explained. The question is what the difference is between the two cases (i.e., initial vs. medial position, (1a) vs. (1b)).

I begin with the observation that what previous authors have called 'medial position' in most cases is more precisely 'after a nasal consonant'.⁷ We know that nasals tend to assimilate to a following obstruent, and my explanation for the difference between initial and medial position lies there. That is, this linking of phonological structure increases resistance to the constraint favoring simplification of the marked cluster /Cʎ/. That is, the intuition is that loss affecting more than one segment is more costly than loss affecting a single segment. That is, /nCʎ/ is more resistant to reduction than simple (word-initial or intervocalic) /Cʎ/ because more segments would be affected.

How may this be formalized? I suggest that this may be handled via the OT instantiation of the Uniformity Condition, whose traditional formulation is given here:

(12) **The Uniformity Condition**

In order to change the feature content of a segment [A], every skeletal slot linked to [A] must satisfy the rule. (Kenstowicz 1994:413)

How may this be captured in a constraint-based approach like OT? I suggest that the effect of this condition may be characterized via constraint conjunction and the formation of a power hierarchy of conjoined constraints with relation to other constraints (see Smolensky 1995). I call this conjoined constraint LINKEDMAX, which is ranked higher than both *COMPLEX and simple MAX (that is, deletion is thwarted because of the linking in [ŋk, mp, mʃ]).⁸

(13) **Fourth proposed historical stage: Retention via Linking.**
(Hispano-Romance, medial position)

'Blocking' of cluster reduction because of nasal assimilation

/*nCk/ retained	LINKEDMAX (NEIGHBORHOOD)	*COMPLEX (ONSET)	MAX
ŋ__k ∨ m__k ∨ mʃ__k ∨	*!		(*)
<input checked="" type="checkbox"/> ŋk ∨ mp ∨ mʃk ∨		*	

I propose that the retention of this cluster via nasal assimilation allows other processes of assimilation to occur, in this case in voicing between the initial consonant and /*k/. This should not be surprising given the analogous devoicing of liquids (and other sonorants) in English (*truck*, *plane*, etc.; Fromkin and Rodman 1988:99), French (*sucre* 'sugar', *pourpre* 'purple', *piéd* 'foot', etc.; Carton 1974: 30-1, 85) and even many varieties of Modern American Spanish, where /tr/ takes on an acoustic similarity to *ch* (= [tʃ]), as in *tronco* 'trunk', often interpreted as *chonco* by the uninitiated (Canfield 1981:7, 13, and *passim*). Furthermore, these changes

often go unnoticed consciously, and so may never be recorded in writing. This assimilation is shown in the following tableau:

(14) Voicing assimilation prevails

/*nC/ > [*ɲc]	LINKEDMAX (NEIGHBORHOOD)	*COMPLEX (ONSET)	MAX	ASSIMILATE
nC		*		*!(vce, PA)
nC		*		*!(PA)
nø	*!		(*)	
<input checked="" type="checkbox"/> ɲc		*		

Examples: MACULA > [*maɲkʎa] > [*maɲcʎa]
 IMPLARE > [*empʎar] > [*eɲcʎir]
 INFLARE > [*iɲʎar] > [*iɲcʎar]

This leads to the next stage in the historical process, that of reinterpretation of [(ɲ)cʎ] as [(n)tʃ] due to their high acoustic similarity (see the spectrograms in Appendix II):

(15) Fifth proposed historical stage: Reinterpretation. High acoustic similarity of [cʎ] to [tʃ]

This acoustic similarity leads to (mis)interpretation of [cʎ] by the listener as [tʃ], and then reanalysis as /tʃ/.⁹ This would be favored by markedness considerations because given the two very different articulations for what is acoustically quite similar, the listener-turned-speaker may choose the simpler of the two. This further optimizes the lexicon by maximizing the harmony of the system (i.e., what is perceived is what is mentally represented, thus reducing the work of the constraints in the grammar.)

Examples: [*maɲcʎa] perceived as [mantʃa], reanalyzed as /mantʃa/
 [*eɲcʎir] = [entʃir] → /entʃir/ *henchir* (*encher* in MPtg.)
 [*iɲcʎar] = [intʃar] → /intʃar/ *hinchar* (*inchar* in MPtg.)

(Additionally, /tʃ/ already exists in Old Spanish (< [jt], e.g., MULTU > H-R, MPtg. *muito* > OSp. *mucho*.)

Thus far I have given an account of the development of initial *lh-* in Spanish, medial [-*lh-*] (*lh*) in Portuguese (which in Spanish then became [x]), and of medial *-ch-* for both Spanish and Portuguese. I have not yet presented an explanation of how Portuguese came to show initial *ch-*.

Recall that I and others have argued that Galician/Portuguese is a more conservative variety of the development of Late Latin. One manifestation of this is that the assimilation of /l/ to /k/ and the extension of /**kl*/ to /**pkl**, /**fl*/ was suggested not to have occurred at the same rate in pre-Old Spanish and Galician/Portuguese. Thus, the simplification of /**Ckl*/ to /k/ did not occur in initial position in Galician/Portuguese because this cluster existed only medially, not initially as in pre-Old Spanish. Once all the /**Ckl*/ clusters are simplified, the constraint *COMPLEX no longer has any candidates that it eliminates, and it fails to play any role in the continued development of these clusters. Given this, I suggest that it is demoted because it is 'inactive'. This would be the sixth historical stage.

- (16) Sixth proposed historical stage: Demotion. Once /*(n)*Ckl*/ is reanalyzed in Hispano-Romance as /(n)tʃ/ there will no longer be any input forms violating the constraint requiring simplification.

At this or a later historical stage, the tendency to assimilate /l/ to /k/ does indeed affect the initial *Cl* clusters of Galician/Portuguese, yielding [**Ckl*] (again, see Wireback 1996 for factors involved in retarded spread of change in Galician/Portuguese). The result is that simplification is no longer the optimal outcome, and more fully assimilated forms prevail.¹⁰

- (17) Creation of *ch-* in Galician/Portuguese

Gal./Ptg. /* <i>Ckl</i> / > [* <i>cʎ</i>]	LINKED MAX	MAX (SON)	MAX (OBS)	*COMPLEX (ONSET)	ASSIMILATE
Cø		*!			
øʎ			*!		
Cʎ				*	*!(vce, PA)
Cʎ̥				*	*!(PA)
<input checked="" type="checkbox"/> cʎ̥				*	

Examples: CLAVE > [*kʎave] > [*cʎave]
 PLUVIA > [*pʎuvja] > [*cʎuvja]
 FLAMMA > [*flama] > [*çʎama]

[cʎ̥] is perceived as [tʃ], and is reanalyzed as /tʃ/ (*chave, chuva, chama*).

This concludes the analysis of the changes of CL, PL, FL from Late to Old Spanish and Galician/Portuguese. In the next sections I elaborate on several of the theoretical points I raised in the previous discussion.

2. The listener as a source of sound change. Given that different vocal tract arrangements may yield similar acoustic speech signals, for the listener there may be articulatory ambiguity. However, the listener aims to pronounce words as nearly as possible in the way she has heard them from others (or thinks she has heard them) (Ohala 1974a,b, 1981, Slobin 1977, Greenlee and Ohala 1980, and for related points, Inkelas 1995, Hale and Reiss 1996, Yip 96).

Given the acoustic similarity of [cʎ] to [tʃ], the listener reconstructs /tʃ/ (incorrectly). This is parallel to the learning systems proposed by Clark and Roberts (1993:301) and Pulleyblank and Turkel (1995a,b,c): Several alternate grammars may adequately account for the input. When this happens, other factors determine the optimal grammar, which in the case described by Pulleyblank and Turkel (1995b) evolves to a more unmarked system. This is 'emergence of the unmarked' (McCarthy and Prince 1993, Smolensky 1996, etc.).

To take the case of the linked clusters, in schematic graphic form we have the following, which shows the passage of phonetic processes to lexicon optimization and the emergence of the unmarked:

(18)		
/Cl/ > [Cʎ]		Does not violate ASSIMILATE(PA) (as much)
/Cʎ/ > [Cʎ]		Does not violate ASSIMILATE(VOICE)
(?) /Cʎ/ > [cʎ]		Does not violate ASSIMILATE(PA)
	/tʃ/ > [tʃ]	Does not violate MARKEDNESS(*ʎ), IDENT

3. Summary and conclusion. To summarize, I explain why Spanish shows different outcome for *Cl* in initial and medial positions, and motivate the 'magic leap' others assumed for the passage of *Cl* to [tʃ]. This was argued to follow from the increased resistance to simplification due to there being linked phonological structure. This was enforced by the OT version of the Uniformity Condition, which then allowed the common processes of voicing and place assimilation to continue. Here the role of the listener is important: there is reinterpretation based on acoustic similarity, markedness considerations and lexicon optimization.

In schematized form, the principal points of the paper are these:

Data: The historical order of changes is summarized below:

palatal assimilation > analogy/allophonic unification > simplification vs. linking (UC) > assimilation and reinterpretation. (The spread of assimilation of #Cɪ to *Cɹ was slower in Galician/Portuguese than in Spanish; when it did occur, the constraint ranking had changed so that reduction was no longer the optimal outcome.)

An additional advance of the proposed analysis is that the process of simplification of Cɹ clusters has now been related to the creation of /tʃ/, which had not connected before.

Issues: Phonetics → phonology → lexicon

The role of the listener (acoustic equivalency, intent to repeat faithfully what heard)

The Uniformity Condition (conjunction and hierarchization), which here prevented simplification from occurring

Lexicon optimization and the emergence of the unmarked ([ɹ] vs. [tʃ], etc.)

Similarity of historical change to child language acquisition and learning systems

Notes

* I'd like to thank Alfonso Morales-Front and Regina Morin for comments and suggestions for improvements. I remain responsible for any deficiencies.

1 In those Latin forms where the consonant and /v/ are not adjacent these two segments came into contact after syncope of the unstressed vowel that separates them. This is exemplified in (5), (7) and (10) below.

2 Here and throughout, a form that has an asterisk before it is not reflected in the written record, but is hypothesized to have existed as an (historically) intermediate stage. Italicized forms show orthography, and words written in small capital letters are Latin forms. Late Latin H = [h].

3 Modern Upper Aragonese is spoken in the upper regions of the province of Aragon, in Spain (near Zaragoza). During the Middle Ages the area where Aragonese was spoken was much greater than that today.

4 This type of 'lifecycle' of a rule is explored in great detail in Janda (1987); I am indebted to Stuart Davis for making me aware of this work.

5 Later, Old Spanish [-ʎ-] was velarized to [x], as in *oreja* 'ear'; Modern Portuguese maintains the final stage cited here. All examples of (7) undergo this change in Old Spanish, and Modern Portuguese maintains the simplified forms without further modification, written *lh* in Portuguese orthography.

6 However, the ranking is opposite that proposed for child language by Gnanadesikan (1995) (Eng. *please* /plɪz/ → [pɪz]).

7 For reasons of space I must omit from consideration those cases where the consonant that precedes the *Cl* cluster is not a nasal. These are treated in Holt (in preparation).

8 This is quite similar in spirit to the constraint NEIGHBORHOOD proposed by Itô and Mester (1996) and earlier work by Joe Pater. This constraint penalizes processes that would affect structure on both sides of a given segment. See Appendix I for two other cases which I suggest may be treated in a manner similar to that proposed here.

9 A very similar proposal is made in Ohala (1974a), where he refutes the purely phonological explanation given by Foley (1973) for the pronunciation in Norwegian of [oʃlo] for *Oslo*. He argues instead for the partial devoicing of [l] by [s]: he then shows that this [ʃ] is acoustically similar to [ʃ], which he believes led to reinterpretation as /ʃ/. For fuller discussion of this and other similar data from Navajo, Algonquian and Itełman, see Holt (in preparation).

10 Alternatively, the Galician/Portuguese reaction to *#Cʎ (or perhaps *#Cʎ) was different, with simplification to [ʎ-] in Spanish but reinterpretation as [tʃ-] in Galician/Portuguese. Since these changes happened in the preliterary period of both Old Spanish and Galician/Portuguese, it is impossible to rule out this alternative, but the proposal given in the text is more in line with the more conservative tendencies attributed to Galician/Portuguese.

Appendix I: Other cases of the 'Uniformity Condition'

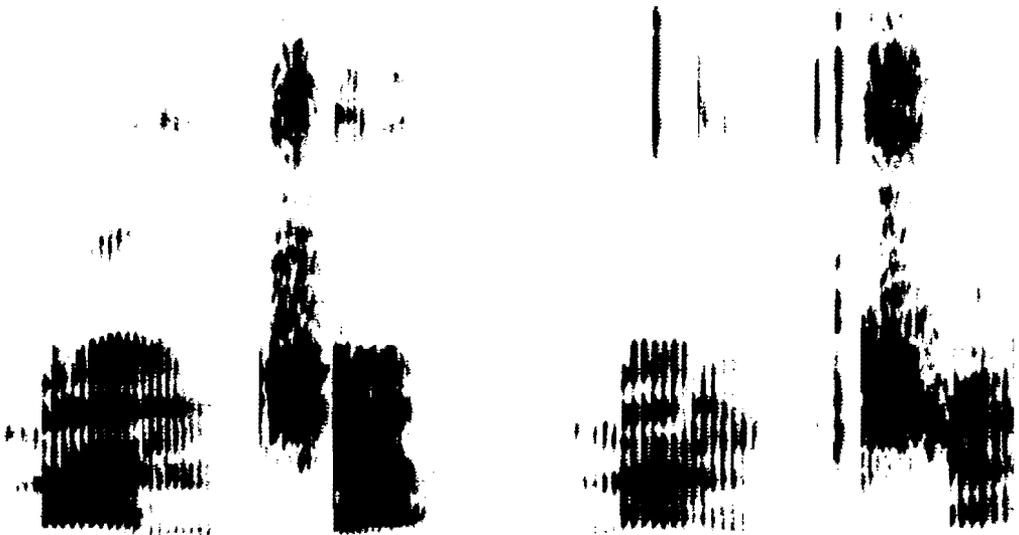
In addition to the case mentioned in the text, I present here two other sets of data which appear to be amenable to a similarly-reformulated Uniformity Condition:

- (a) Loss of stop element of Proto Indo-European *g^w is blocked when a nasal consonant precedes it: e.g. PIE *g^wiōu > CL VIVUS 'living' vs. *d^hg^wā > LINGUA 'tongue' (Ohala 1981).

- (b) Vocalization-cum-palatalization in Old Spanish: e.g. OCTO 'eight' > [*oxto] > [ojto] > *ocho*
 Thwarted when more than one consonant would be affected: e.g.,
 VULTURE > *buitre*, not [*butr'e] (Penny 1991) (also FRAXINU,
 SEX, PECTINARE, PIGNORA, etc.)

In each case a conjoined constraint dominates a structural constraint (simplification or palatalization, respectively), which in turn dominates the relevant simple constraint.

Appendix II: High acoustic similarity between [tʃ] and [cʃ]



[mantʃa]

[maɲcʃa]

References

- Bourciez, Edouard. 1967. *Eléments de linguistique romane*. 5th. ed. Paris: Klincksieck.
- Canfield, D. Lincoln. 1981. *Spanish Pronunciation in the Americas*. Chicago: University of Chicago Press.
- Carton, Fernand. 1974. *Introduction a la phonétique du français*. Paris: Bordans.
- Clark, Robin and Ian Roberts. 1993. A computational model of language learnability and language change. *Linguistic Inquiry* 24.299-345.
- Foley, James. 1973. Assimilation of phonological strength in Germanic. In Stephen R. Anderson and Paul Kiparsky, eds., *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston. 51-58.
- Fromkin, Victoria and Robert Rodman. 1988. *An Introduction to Language*. New York: Holt, Rinehart and Winston. 4th. ed.
- Gnanadesikan, Amalia E. 1995. Markedness and faithfulness constraints in child phonology. Ms. Rutgers University. Rutgers Optimality Archive.
- Greenlee, Mel and John J. Ohala. 1980. Phonetically motivated parallels between child phonology and historical sound change. *Language Sciences* 2.283-304.
- Hale, Mark and Charles Reiss. 1996. The comprehension/production dilemma in child language: A response to Smolensky. Ms. Concordia University, Montréal. Rutgers Optimality Archive.
- Hargus, Sharon. 1995. The first person plural subject prefix in Babine/Witsuwit'en. Ms. University of Washington. Rutgers Optimality Archive.
- Holt, D. Eric. In preparation. *Constraint Interaction in the Development of Syllable Structure from Latin to Spanish and Portuguese*. PhD Dissertation. Georgetown University.
- Inkelas, Sharon. 1995. The consequences of Optimization for Underspecification. *Proceedings of NELS* 25. 287-302.
- Itô, Junko and Armin Mester. 1996. Structural economy and OCP interactions in local domains. Paper presented at the Western Conference on Linguistics (WECOL), University of California, Santa Cruz, October 25-27, 1996.
- Janda, Richard David. 1987. *On the Motivation for an Evolutionary Typology of Sound-Structural Rules*. PhD Dissertation. UCLA.
- Janson, Tore. 1977. Latin vowel reduction and the reality of phonological rules. *Studia Linguistica* 31.1-17.
- Jonasson, J. 1971. Perceptual similarity and articulatory reinterpretation as a source of phonological innovation. *Speech Transmission Laboratory (Stockholm), Quarterly Progress and Status Report* 1.30-41.

- Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Cambridge, Mass.: Blackwell.
- Lloyd, Paul M. 1987. *From Latin to Spanish*. Philadelphia: American Philosophical Society. Vol. 173.
- Malkiel, Yakov. 1963-4. The interlocking of narrow sound change, broad phonological pattern, level of transmission, areal configuration, sound symbolism: Diachronic studies in the Hispano-Romance consonant clusters *ct-*, *fl-*, *pt-*. *Archivum Linguisticum* 15.144-173, 16.1-33.
- McCarthy, John J. and Alan Prince. 1993. *Prosodic Morphology: Constraint Interaction and Satisfaction*. Ms. University of Massachusetts, Amherst and Rutgers University.
- Niyogi, Pradip and Robert C. Berwick. 1995. The logical problem of language change. AI Memo 1516. CBCL Paper 115, MIT.
- Ohala, John J. 1974a. Phonetic explanation in phonology. *Proceedings of CLS*. 251-274.
- Ohala, John J. 1974b. Experimental historical phonology. In J.M. Naderson and C. Jones, eds., *Historical Linguistics II. Theory and description in phonology*. *Proceedings of the First International Conference on Historical Linguistics, Edinburgh 2nd-7th September 1973*. New York: Elsevier. 353-389.
- Ohala, John J. 1980. The application of phonological universals to speech pathology. In N.J. Lass, ed., *Speech and Language: Advances in Basic Research and Practice*. Vol. 3. New York: Academic Press. 75-97.
- Ohala, John J. 1981. The listener as a source of sound change. In M.F. Miller, et al., eds., *Proceedings of CLS*.
- Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage, ed., *The Production of Speech*. New York: Springer.
- Penny, Ralph. 1991. *A History of the Spanish Language*. Cambridge: Cambridge University Press.
- Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms. Rutgers University and University of Colorado, Boulder.
- Pulleyblank, Douglas and William J. Turkel. 1995a. The logical problem of language acquisition in Optimality Theory. To appear in *Proceedings of the Workshop on Optimality in Syntax—Is the Best Good Enough?* MIT Press and MIT Working Papers in Linguistics.
- Pulleyblank, Douglas and William J. Turkel. 1995b. Optimality Theory and learning algorithms: The representation of recurrent featural asymmetries. Paper presented at *Current Trends in Phonology (CTIP)*, Abbaye de Royaumont, June 1995.

- Pulleyblank, Douglas and William J. Turkel. 1995c. Gradient ascent learning in an Optimality Theoretic Framework: Violable constraints and the treatment of local maxima. Paper presented at Workshop on Derivations and Constraints in Phonology, Essex, August 1995.
- Repetti, Lori and Edward F. Tuttle. 1987. The evolution of Latin *PL*, *BL*, *FL*, and *CL*, *GL* in Western Romance. *Studi Mediolatini e Volgari* 22.53-115.
- Slobin, Dan I. 1977. Language change in childhood and in history. In John Macnamara, ed., *Language Learning and Thought*. New York: Academic Press. 185-214.
- Smolensky, Paul. 1995. On the internal structure of the constraint component *Con* of UG. Handout of talk at Johns Hopkins University. Spring 1995.
- Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27.
- Tuttle, Edward F. 1975. The development of *pl*, *bl* and *fl* in Italo-Romance: Distinctive features and geolinguistic patterns. *Revue de Linguistique Romane* 39.400-431.
- Williams, Edwin B. 1938. *From Latin to Portuguese: Historical Phonology and Morphology of the Portuguese Language*. Philadelphia: University of Pennsylvania Press.
- Wireback, Kenneth J. 1996. On the palatalization of /k/, /p/, and /f/ in Hispano-Romance. Paper presented at the 49th Kentucky Foreign Language Conference, University of Kentucky, April 19, 1996.

Linearization of Verb Clusters in West Germanic

Andreas Kathol
UC Berkeley

1 Introduction

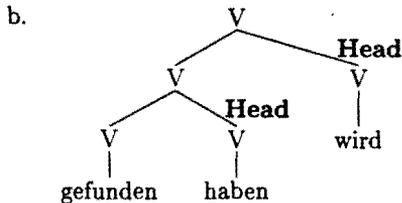
One often debated issue in German syntax has been whether clause-final sequences of verbs such as *gefunden haben wird* in (1a) involve a constituent structure comparable to that commonly assumed for the corresponding English example in (1b).¹

- (1) a. daß Peter das Buch **gefunden haben wird**.
 that Peter the book found-PSP have-INF will-FIN
 ‘that Peter will have found the book.’
 b. that Peter [will [have [found [the book]]]].

A venerable tradition going back to Evers (1975) and earlier holds that German is indeed similar to English in that a main verb forms a constituent with its complements at an underlying level, but the latter are subsequently moved to higher positions, leaving behind a purely verbal string.

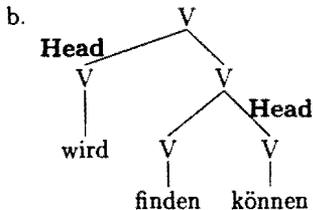
An alternative line of thought has emerged in recent years which instead regards verbal sequences as **single** underived syntactic entities, to be referred to here as **verbal complexes**. While this talk will focus on approaches along these lines advanced in Head-Driven Phrase Structure Grammar (Pollard & Sag 1994), there have also been proposals within transformational frameworks, most notably Bierwisch (1990) and Haider (1993). In all these works, the verbal grouping indicated in (2a) directly results from the combinatorics of the verbal heads without recourse to underlying “phrasal” constituents. Thus, German head-final verb clusters involve sequences of binary combinations resulting in a left-branching structure, as shown in (2b).

- (2) a. daß Peter das Buch [gefunden haben wird]



Evidence for this analysis has been adduced for instance by Hinrichs & Nakazawa (1989), who point out that the posited constituents are precisely those that are affected by the order alternation known as *Oberfeldumstellung* or *Aux Flip*, demonstrated in (3a). Thus, while governors in German usually follow, certain environments require or allow the tense auxiliaries *haben* and *werden* to precede their verbal complement. The structure for such orders under the verbal complex analysis is outlined in (3b):

- (3) a. daß Peter das Buch **wird** finden können/ **hat** finden können.
 that Peter the book will find can has find can
 'that Peter will be able to find the book.'
 'that Peter has been able to find the book.'



Moreover, precisely those constituents postulated by a left-branching analysis also surface in V2 fronting constructions as shown in (4a). Smaller constituents can also be affected, including single governed verbs, as in (4b). Under this view these are simply single-element verbal subcomplexes.

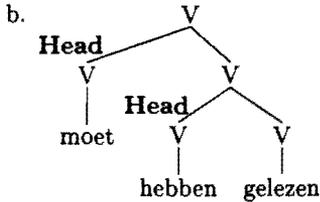
- (4) a. [**Finden können**] wird Peter das Buch.
 find can will-FIN Peter-NOM the book-ACC
- b. [**Finden**] wird Peter das Buch können.
 find will-FIN Peter-NOM the book-ACC can

1.1 Dutch

Despite its general close similarity with German, one area in which Dutch exhibits a striking difference is in the organization of the verbal complex.

With head-initial structures being possible in a larger class of environments, we standardly get orderings that constitute the mirror image of the German cases. In (5a) an example is given, together with its right branching analysis in (5b).

- (5) a. dat Jan dit boek moet₁ hebben₂ gelezen₃.
 that Jan this book must-FIN have-INF read-INF
 'that Jan must have read the book.'



However, Dutch also allows German-style head-final orderings with certain types of verbal governors. As shown in (6), tense auxiliaries such as *hebben* can optionally follow their verbal complement.

- (6) dat Joop de krant heeft gelezen/ gelezen heeft.
 that Joop the newspaper has read read has
 'that Joop read the newspaper.'

1.2 Challenges

The view of the verb cluster in West Germanic just presented embodies the assumption that all possible structures arise from reorderings among subconstituents of the verbal complex. Moreover, since the construction of verbal complexes involves binary trees, there is a prediction that a verb should always occur as preceding or following the subcomplex it governs.

This latter assumption faces a severe challenge from constructions in both languages. In Dutch, the head-final order permitted for tense auxiliaries in (6) is not limited to finite cases. Thus, infinitival *hebben* can follow the participial main verb *gelezen*, as in (7b), giving rise to a structure isomorphic to the German *Oberfeldumstellung* case in (3). However, in many dialects, this order is decidedly dispreferred in comparison to the order in (7c). Here, *moet* occurs in between the elements of the governed subcomplex *gelezen hebben*.

- (7) a. dat Jan dit boek moet₁ hebben₂ gelezen₃.
 that Jan this book must-FIN have-INF read-INF
 'that Jan must have read the book.'
- b. dat Jan dit boek moet₁ gelezen₃ hebben₂.
 that Jan this book must-FIN read-PSP have-INF

- c. dat Jan dit boek gelezen₃ moet₁ hebben₂.
 that Jan this book read-PSP must-FIN have-INF

In a similar fashion, Meurers (1994) has argued that parallel perturbations of the canonical order in German are more common than have generally been thought. Attested examples in which a finite governor occurs inside the subcomplex it selects are given in (8):

- (8) a. daß er das Examen bestehen₃ wird₁/hat₁ können₂.
 that he the exam pass will/has can
 'that he will be/has been able to pass the the exam.'
- b. zu dem Zeitpunkt an dem ich mich entscheiden₃ hätte₁ müssen₂
 at the point at which I me decide had must
 'at the point at which I should have made a decision'

The most uniform analysis assigns the same selectional and constituency relations among the cluster elements in all three Dutch examples in (7). This means that all three sentences in (7) involve the partial verb cluster *gelezen hebben*, even when it does not occur as a contiguous string, as in the case of (7c). In an analogous fashion, the German examples in (8) involve the discontinuous realization of the verbal subcomplexes *bestehen können* and *entscheiden müssen*, respectively. However, rather than assume some movement-based restructuring operation that derives (7c) and (8a,b) from some underlying canonical structure, I propose to treat the discontinuity at face value and as the result of the linearization conditions imposed on elements of the verb cluster. This will be done by adopting a variant of HPSG which assumes a somewhat looser correlation between order and constituency than is possible with strictly phrase structure-based models.

2 Argument Composition

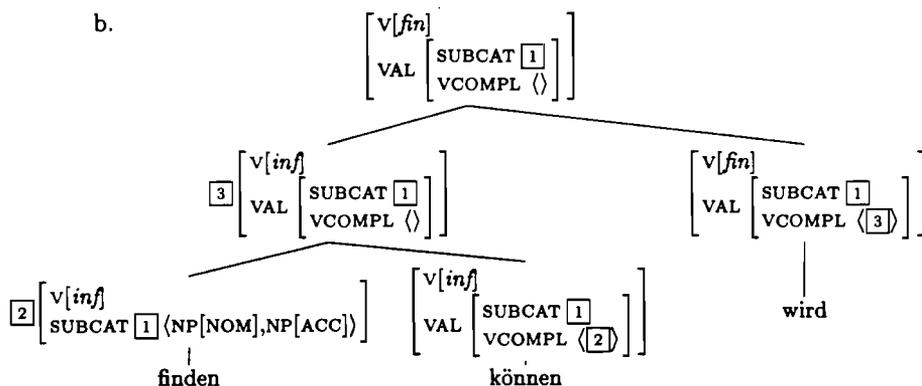
Before going into the details of the analysis, we need to briefly review Hinrichs & Nakazawa's (1989) analysis of verbal complexes in terms of "argument composition". Argument composition can be viewed as a form of generalized raising and has the result of merging the valence properties of the participating verbs. Thus, while the analysis of raising structures in English in nonderivational theories such as HPSG involves structure-sharing of the understood subject of a VP complement with that of the raising verb's subject, argument composition raises all arguments of an embedded verb. The latter then become part of the higher predicate's valence, as indicated in (9a). The valence of the entire complex then consists of the list of arguments "attracted" in step-wise fashion from lower predicates.

Among proponents of argument composition, it is commonly assumed that the same valence attribute selecting phrasal arguments is also involved in the selection of the verbal complement, viz. SUBCAT or SUBJ and COMPS. The verbal complement itself is added as a valence element of the raiser by appending it (notated as “o”) to the list of raised arguments, as shown in (9a). However, as I have argued elsewhere (Kathol 1995a, Kathol Forthcoming), there are both technical and empirical reasons to delegate the selection of verbal elements to a different attribute, here called VCOMPL. If we adopt this differentiation in valence attributes then the description of an predicate attracting all the arguments of its verbal complement is as outlined in (9b).

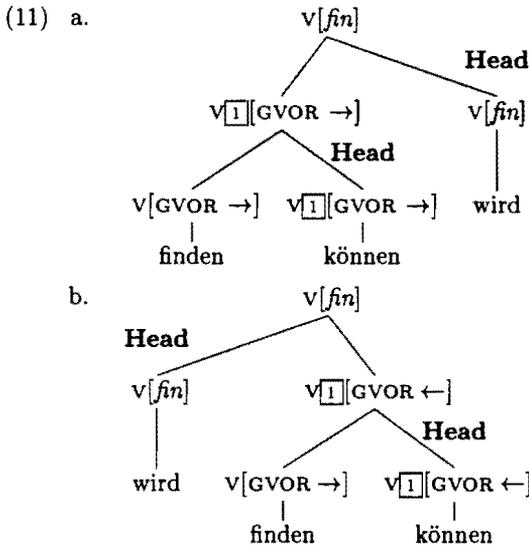
- (9) a. $[\dots | \text{SUBCAT } \boxed{1} \circ \langle v [\dots | \text{SUBCAT } \boxed{1}] \rangle]$
 b. $[\dots | \text{VALENCE } \left[\begin{array}{l} \text{VCOMPL } \langle v [\dots | \text{SUBCAT } \boxed{1}] \rangle \\ \text{SUBCAT } \boxed{1} \end{array} \right]]$

In (10b) I show how the combination of elements in the verbal cluster interacts so as to allow for the entire complex to have the valence of the embedded main verb.

- (10) a. daß Peter das Buch finden können wird.
 that Peter the book find can will
 ‘that Peter will be able to find the book.’



Focussing on order, I will adopt Hinrichs & Nakazawa's (1990) idea that a nonfinite verb constrains the possible relative position of its direct governor. This is done here by means of the head attribute GVOR.² For now we will distinguish two values, "→" and "←", indicating placement of a higher governor to the right or left, respectively. Since *können* permits both orders, we either get the canonical order as in (11a) or the *Oberfeldumstellung* order shown in (11b):



The distinction in order depending on the GVOR value is straightforwardly captured by the linear precedence constraints in (12):

- (12) a. (verbal complement) (precedes) (governor)
 $[\text{SYNSEM } \boxed{1} [\dots | \text{GVOR } \rightarrow]] < [\dots | \text{VCOMPL } \langle \boxed{1} \rangle]$
- b. (governor) (precedes) (verbal complement)
 $[\dots | \text{VCOMPL } \langle \boxed{1} \rangle] < [\text{SYNSEM } \boxed{1} [\dots | \text{GVOR } \leftarrow]]$

As a result, German and Dutch can be treated as on a par with respect to the constituent structure of verbal complexes, while differing in the distribution of precedence indicators among lexical classes of verbs. For instance, nonfinite main verbs in Dutch are underspecified in terms of their value for GVOR while German main verbs obligatorily require placement to the right, as shown in (13).³

(13) Classification of nonfinite main verbs

- a. Dutch
 $[\dots | \text{HEAD} | \text{GVOR } \textit{dir}]$
- b. German
 $[\dots | \text{HEAD} | \text{GVOR } \rightarrow]$

This classification correctly allows for the order variability with Dutch main verbs seen earlier in (6). By contrast, German main verbs must always precede any direct governor, as is demonstrated in (14).

- (14) *daß Lisa die Zeitung hat gelesen.
 that Lisa the newspaper has read

3 Discontinuous verbal complexes

The central property that unites all the problematic ordering possibilities in Dutch and German seen earlier is that the crucial determinant for order is not the selected verbal complex itself, but its **lexical head**. Thus, returning to the possible orders in (7) above, repeated in (15a-c), in all the grammatical cases, the governor *moet* precedes the head of the selected complex, *hebben*. As soon as the order is reversed, i.e., as soon as *moet* follows the head of the governed complex *hebben*, the result becomes ungrammatical, as shown in (15d,e).

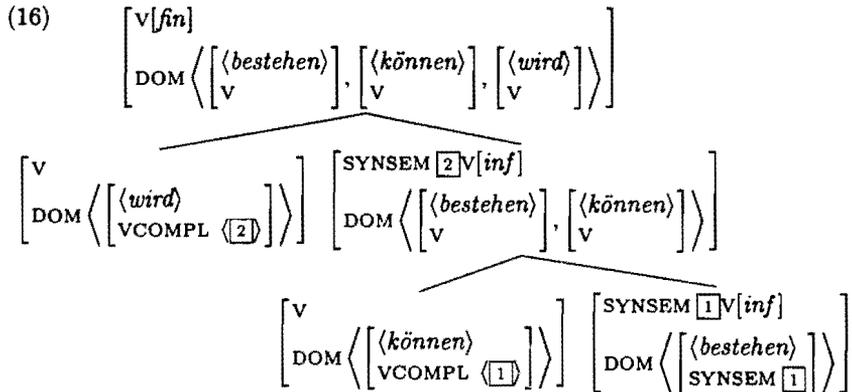
- (15) a. dat Jan dit boek moet₁ hebben₂ gelezen₃.
 that Jan this book must-FIN have-INF read-INF
 'that Jan must have read the book.'
- b. dat Jan dit boek moet₁ gelezen₃ hebben₂.
 that Jan this book must-FIN read-PSP have-INF
- c. dat Jan dit boek gelezen₃ moet₁ hebben₂.
 that Jan this book read-PSP must-FIN have-INF
- d.*dat Jan dit boek gelezen₃ hebben₂ moet₁.
 that Jan this book read-PSP have-INF must-FIN
- e.*dat Jan dit boek hebben₂ moet₁ gelezen₃.
 that Jan this book have-INF must-FIN read-PSP

What is therefore needed is a way to allow for the component parts of certain syntactic elements to be "visible" for the placement options of any higher governor.

This is precisely what can be achieved by the adoption of **order domains** which allow us to extend the range within which syntactic elements interact in their ordering properties beyond the scope of ordinary constituency. In the normal mode of combination, the internal components of, say, an NP argument will be rendered opaque for interleaving by other syntactic elements. By contrast, the current proposal advocates a transparent mode of combination in the case of verbal complexes. Hence, the idea is somewhat reminiscent of the notion of **liberation** proposed by Pullum and Zwicky for GPSG (Pullum 1982, Zwicky 1986). However, rather than operate on rules, order domains allow one to leave the combinatorial system itself uniform and locate the effect of different modes of combination entirely in the linearization component.

As a concrete example, consider the domain construction associated with a "canonical" German verb cluster, as in (16). Here, order domains are given

as the value of DOM, taking a totally ordered list of phonology-category pairs as its value. The phonology of the entire sign is simply the phonology values of all domain elements, strung together in the same order.



When *bestehen* and *können* are combined, each contributes only one domain element, hence by necessity, the two will occur adjacent to each other, no matter what order is chosen. However, when the resulting subcomplex is combined with *wird*, there are now three placement options: before *bestehen können*, between *bestehen* and *können*, and following *bestehen können*. In the case of canonical orderings, only the last is grammatical, as required by the specification [GVOR →] on the subcomplex as inherited from *können*. Yet, that subcomplex can no longer be referred to as a separate element within the order domain. To accommodate the transparency effect, we modify the LP constraint in (12) by making reference not to the entire governed complex, but only its **head**. The LP constraints in (17) achieve the desired effect by structure sharing in the HEAD values and hence supersede the earlier formulation in (12):

- (17) a. (head of verbal complement) (precedes) (governor)

$$\left[\dots | \text{HEAD } \boxed{1} [\text{GVOR } \rightarrow] \right] \prec \left[v[\text{VCOMPL } \langle \text{HEAD } \boxed{1} \rangle] \right]$$
- b. (governor) (precedes) (head of verbal complement)

$$\left[v[\text{VCOMPL } \langle \text{HEAD } \boxed{1} \rangle] \right] \prec \left[\dots | \text{HEAD } \boxed{1} [\text{GVOR } \leftarrow] \right]$$

If the governor has to follow the head of the governed verbal complex, it necessarily has to follow *all* elements of the verbal complex, hence reference to the head in a head-final structure has the same effect as reference to the entire cluster as before in (12a).

Now let us consider the case in which *können*'s GVOR value is instantiated as ←. Any governor will have to precede it, but there is *prima facie* no

requirement that precedence has to be *immediate* and therefore the lowest governed verb, *bestehen*, may intervene between the two. This is possible so long as *bestehen*'s linear relation to its governor is in accordance with its own GVOR value. As a German main verb, it requires the governor to follow— as a result, both precedence requirements are satisfied in the order domain in (18), where arrows relate verbs and the heads of the verbal complex that they govern:

$$(18) \left[\text{DOM} \left\langle \left[\langle \text{wird} \rangle \right], \left[\begin{array}{c} \langle \text{bestehen} \rangle \\ \text{GVOR} \rightarrow \end{array} \right], \left[\begin{array}{c} \langle \text{können} \rangle \\ \text{GVOR} \leftarrow \end{array} \right] \right\rangle \right]$$

However, this is not the only solution of the linearization constraints among the elements of the pairs *bestehen*–*können* and *wird*–*können*. A second possibility consistent with the LP requirements is to place the highest governor *wird* immediately before the head of the governed complex, *können*. The latter in turn only has to follow the dependent verb *bestehen*, but not immediately. As a result, we obtain a situation in which the governed subcomplex *bestehen können* is linearized in a discontinuous fashion, as shown in (19):

$$(19) \left[\text{DOM} \left\langle \left[\begin{array}{c} \langle \text{bestehen} \rangle \\ \text{GVOR} \rightarrow \end{array} \right], \left[\langle \text{wird} \rangle \right], \left[\begin{array}{c} \langle \text{können} \rangle \\ \text{GVOR} \leftarrow \end{array} \right] \right\rangle \right]$$

3.1 Immediate precedence

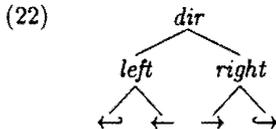
While the evidence presented here is highly suggestive of a looser relationship between constituency and linear order, there nevertheless are environments that do not permit intrusion effects. One striking difference can be observed in the behavior of separable prefixes in Dutch vs. German. In Dutch, a separable prefix such as *aan* may “float” to earlier positions away from its base verb, here *spreken*. Such a discontinuous realization of a particle-verb combination is shown in (20b).

- (20) a. dat Jan Marie heeft aangesproken.
 that Jan Marie has PREF.spoken
 ‘that Jan addressed Marie.’
- b. dat Jan Marie aan heeft gesproken.
 that Jan Marie PREF has spoken

By contrast, most German dialects are much more resistant against “floating” prefixes, as can be seen from the ungrammaticality in (21b):

- (21) a. daß Hans Maria wird ansprechen wollen.
 that Hans Maria will PREF.speak want
 'that Hans will have wanted to address Maria.'
- b. *daß Hans Maria **an** wird **sprechen** wollen.
 that Hans Maria PREF will speak want

What this suggests is that different dialects not only vary in terms of the **linear** constraints between verbal governors and dependent elements, but also in terms of **adjacency conditions**. With respect to our formalization using GOVR, this means that in addition to the linear precedence encoded via \rightarrow and \leftarrow , we also have **immediate precedence** to the *right* or *left*. Let us therefore assume the additional GOVR values \leftrightarrow and \leftarrow . The whole range of precedence indicators is then the one given by the hierarchy for GOVR values in (22):



Values requiring adjacency are referenced by the immediate precedence constraints in (23).

- (23) a. (head of verbal compl.) (imm'ly precedes) (governor)
 $[\dots | \text{HEAD } \boxed{1} | \text{GOVR } \leftrightarrow] \quad \leftarrow \quad [\text{V} [\text{VCOMP } \langle [\text{HEAD } \boxed{1}] \rangle]]$
- b. (governor) (imm'ly precedes) (head of verbal compl.)
 $[\text{V} [\text{VCOMP } \langle [\text{HEAD } \boxed{1}] \rangle] \quad \leftarrow \quad [\dots | \text{HEAD } \boxed{1} | \text{GOVR } \leftarrow]]$

Since separable prefixes distributionally behave like governed nonfinite verbs as part of the verbal complex, their placement with respect to their base verb can be subsumed under the present analysis using the GOVR attribute. As shown in (24), Dutch prefix verbs require dependent prefixes to precede, while German imposes **immediate** precedence.

(24) Valence of particle verbs

- a. German
 $[\dots | \text{VALENCE } [\text{VCOMPL } \langle \text{P} [\text{GOVR } \leftrightarrow] \rangle]]]$
- b. Dutch
 $[\dots | \text{VALENCE } [\text{VCOMPL } \langle \text{P} [\text{GOVR } \rightarrow] \rangle]]]$

However, it is important to keep in mind that any of the LP constraints proposed here only require precedence and/or adjacency if the two elements actually belong to the verb cluster. None of the cluster-specific placement requirements carry over to the case where the finite verb occurs in clause-initial or second position in root environments, illustrated in (25):

- (25) a. Jan **spreekt** Marie **aan**.
 Jan speaks Marie PREF
 'Jan addresses Marie.'
- b. Hans **spricht** Maria **an**.
 Hans speaks Maria PREF
 'Hans addresses Maria.'

As I have shown elsewhere (Kathol 1995a, Kathol 1995b), the suspension of any precedence constraints on governors in root clause positions can be captured in a natural way by partitioning the domain of the entire clause into positional classes, outlined in (26a) via the class indicators *comp*, *mittelfeld*, and *verb cluster*.

- (26) a. $\left[\text{DOM} \left\langle \dots, \left[\begin{array}{c} \text{"comp"} \\ \langle \text{spreekt} \rangle \\ \text{V[FIN]} \end{array} \right], \left[\begin{array}{c} \text{"mittelfeld"} \\ \langle \text{Marie} \rangle \\ \text{NP} \end{array} \right], \left[\begin{array}{c} \text{"verb cluster"} \\ \langle \text{aan} \rangle \\ \text{P} \end{array} \right] \right\rangle \right]$
- b. $\left[\text{DOM} \left\langle \dots, \left[\begin{array}{c} \text{"mittelfeld"} \\ \langle \text{Marie} \rangle \\ \text{NP} \end{array} \right], \left[\begin{array}{c} \text{"verb cluster"} \\ \langle \text{aan} \rangle \\ \text{P} \end{array} \right] \left[\begin{array}{c} \text{"verb cluster"} \\ \langle \text{spreekt} \rangle \\ \text{V[FIN]} \end{array} \right] \right\rangle \right]$

Then the scope of all ordering statements pertaining to verbal complex elements has to be understood as restricted to that that positional class, viz. *verb cluster*. As can further be shown, the required distinction in positional classes is sufficient to account for the different placement options for finite verbs without any need to posit head movement operations to derive root placement from some underlying position. As illustrated in (26), the different placement options of finite verbs simply correspond to assignments to different positional classes, i.e., *comp* vs. *verb cluster*.

4 Conclusion

In conclusion, intrusion effects appear at first to challenge the hypothesis that all West Germanic dialects are basically alike in terms of the government and constituency relations holding within the verb cluster. Once it is realized, however, that the ordering relations in question should be thought of as relations among heads then a somewhat different pattern emerges that contains the seemingly well-behaved structures as a special case. In order to implement this idea, it is necessary to dissociate the determination of order from immediate syntactic constituency.

If the foregoing is on the right track, it suggests that much, if not all, of the parametric variation in West Germanic verb clusters can be reduced to lexical variation involving requirements of precedence and adjacency.

Endnotes

¹I would like to thank the participants of the Wecol conference for useful discussion. Thanks also go to Gosse Bouma, Jack Hoeksema, John Nerbonne, and Gertjan van Noord for discussions of earlier version of this work. The usual disclaimers apply.

²This attribute thus supersedes Hinrichs and Nakazawa's binary-valued attribute FLIP.

³Here, *dir* is a sort subsuming the subtypes ← and →; cf. also (22) below.

References

- BIERWISCH, MANFRED. 1990. Verb cluster formation as a morphological process. In *Yearbook of Morphology*, ed. by Gert Booij & Jaap van Marle, 173-199. Dordrecht: Foris.
- EVERS, ARNOLD. 1975. *The Transformational Cycle in Dutch and German*. University of Utrecht dissertation.
- HAIDER, HUBERT. 1993. *Deutsche Syntax - generativ. Vorstudien zur Theorie einer projektiven Grammatik*. Tübingen: Gunter Narr.
- HINRICHS, ERHARD, & TSUNeko NAKAZAWA. 1989. Flipped out: Aux in German. In *Papers from the 25th Meeting*, 193-202, Chicago. Chicago Linguistic Society.
- , & ——. 1990. Subcategorization and VP Structure in German. In *Proceedings of the Third Symposium on Germanic Linguistics*, ed. by Shaun Hughes & Joe Salmons. Benjamins, Amsterdam.
- KATHOL, ANDREAS. 1995a. *Linearization-Based German Syntax*. Ohio State University dissertation.
- . 1995b. Verb-‘Movement’ in German and Topological Fields. In *Papers from the 31st Regional Meeting, Chicago Linguistic Society*, ed. by Audra Dainora, Rachel Hemphill, Barbara Luka, Barbara Need, & Sheri Pargman, volume 31, 231-245, Chicago, Illinois. CLS, CLS.
- . Forthcoming. Constituency and Linearization of Verbal Complexes. In *Complex Predicates in Nonderivational Syntax*, ed. by Erhard Hinrichs, Andreas Kathol, & Tsuneko Nakazawa, Syntax and Semantics. New York: Academic Press.
- MEURERS, WALT DETMAR. 1994. A modified view of the German verbal complex. Presentation given at the 1994 HPSG workshop in Heidelberg.

- POLLARD, CARL J., & IVAN A. SAG. 1994. *Head-Driven Phrase Structure Grammar*. CSLI Publications and University of Chicago Press.
- PULLUM, GEOFFREY. 1982. Free Word Order and Phrase Structure Rules. In *Proceedings of the 12th Annual Meeting of the Northeast Linguistic Society*, ed. by James Pustejovsky & Peter Sells, 209–20. Amherst: GLSA.
- ZWICKY, ARNOLD. 1986. Concatenation and Liberation. In *Papers from the 22nd Regional Meeting, Chicago Linguistic Society*, 65–74.

Emergence of the markedness constraint No-Lar and its split behavior in the reduplicative domain*

Soohee June Kim
soohee@u.washington.edu

1. Introduction

This paper investigates the behavior of two reduplicants CV and CVC in Korean with respect to laryngeal feature copying. The CVC reduplicant in Korean copies laryngeal features of the base, whereas the CV reduplicant does not. Since No-Lar(yngeal), which bans marked laryngeal features (Lombardi 1995), can be considered a markedness constraint, the Korean case poses a problem to the position that effects of markedness constraints are visible indiscriminately in the reduplicative domain (cf. emergence of the unmarked, McCarthy and Prince 1994b). Resorting to McCarthy and Prince's (1994a) claim that the RED takes on the characteristics of affix or stem of normal phonology, I will argue that the Korean CV reduplication is an instantiation of the reduplicative template (RED) classified as affix in the lexicon, whereas the CVC RED as stem. This conclusion follows from the fact that the CV RED aims for less marked structure with respect to features just like regular affixes and that the CVC RED cares more for the identity relationship (Ident-BR) just like the forms in normal phonology (Ident-IO).

Following the discussion of the behavior of laryngeal features in normal phonology in section 2, a partial reduplication process (base+CV reduplicant) in Korean is analyzed in section 3 in the Optimality framework as a case of emergence of the unmarked (cf. McCarthy and Prince 1994b). In the next section, another partial reduplication process in Korean (base+CVC reduplicant) is presented which preserves the input laryngeal features in the reduplicant. This is a puzzle, since not all reduplicants demonstrate emergence of the unmarked as expected. In section 5, I propose to solve the puzzle by identifying the two types of reduplicant with the respective affix and stem in normal phonology. Potential problems and theoretical implications are discussed in section 6.

2. Normal phonology and feature specification of Lar in Korean

Before the analysis of the laryngeal feature copying in CV and CVC reduplication, a look at laryngeal specification of the obstruents in the normal (non-

reduplicative) phonology of Korean is necessary. As is well known, obstruents in Korean have a three-way distinction: the plain, the aspirated, and the tense¹. I will use bilabial stops for an illustration. As seen in (1) below, all three types are allowed in the syllable-initial position on the surface²:

- (1) plain: pul 'fire'
 aspirated: p^hul 'grass'
 tense: p'ul 'horn'

Following the general practice, I will adopt the privative feature specification and use constricted glottis [CG] and spread glottis [SG] to specify laryngeal feature of the Korean tense and the aspirated obstruents respectively (cf. see note 3 for reasons to reject the binary specification):

- (2) [] for p in pul
 [SG] for p^h in p^hul
 [CG] for p' in p'ul

Apart from the implicational universal based on the cross-linguistic generalization that glottalized or aspirated obstruents are rarer than their plain counterparts (Maddison 1984) or the intuition that less structure means less marked, the unmarked status of the plain obstruents with respect to the laryngeal node is evidenced in Korean by two phonological phenomena. First, examples in (3) below show that laryngeally marked obstruents lose their laryngeal features and get neutralized into the plain ones in the syllable-final position (see Kim 1974 or You-Cho 1988 for detailed discussion of the issue):

	<u>UR</u>	-->	<u>Surface</u>	
(3)	/kuk/		[kuk]	(no change)
	soup			
	/pu.ək ^h /	-->	[pu.ək]	(aspiration loss)
	kitchen			
	/sək'-tə/	-->	[sək-t'a]	(tensing loss)
	mix-mood marker			

The bold-faced underlined codas above have lost their underlying laryngeal specification, which can be understood as a manifestation of unmarked segments in

the syllable-final position (Iverson and Kim 1987). Similar phenomenon has been analyzed in other languages as delinking of specific laryngeal features, resulting in empty laryngeal node (e.g. Thai by Clements 1985).

Second, only plain obstruents are voiced intervocalically. If the assumption is correct that the plain obstruents are empty in the laryngeal node, that is, underlyingly unmarked, intervocalic assimilatory voicing of the plain consonant in (4) below can be easily explained as spreading of the feature [voice] to the empty laryngeal node:

(4)	/i tal / -->	[i dal]	(voicing)	'this moon'
	/i tʰoŋ / -->	[i tʰoŋ]	cf. *[i doŋ]	'this container'
	/i t'aŋ / -->	[i t'aŋ]	cf. *[i daŋ]	'this land'

If Korean allows only one feature under the laryngeal node, why intervocalic voicing is not applied to the laryngeally marked obstruents is easily accounted for. The underlying laryngeal features [SG] and [CG] already occupy the laryngeal node and will block the docking of the voice feature³ (For a more generalized account for the unmarked status of plain voiceless laryngeal obstruents, see Lombardi 1991).

We can now safely say that the plain obstruents in Korean (C) is unmarked with respect to the laryngeal feature whereas the aspirated (C^h) and the tense (C') are marked.

As said already, all three obstruent types discussed in Korean (C^h, C', and C) appear freely in the onset position. That is, the input laryngeal features faithfully surface in the output in the onset position as evidenced in a word like [p'ul] "horn" (cf. *[pul] for "horn"). Adopting the correspondence theory, I will use two constraints to explain the feature faithfulness, namely Ident-IO(Lar) and No-Lar. Ident-IO(Lar) will govern the input-output feature faithfulness, specifically laryngeal features, and No-Lar, as a markedness constraint, will militate against marked features, namely laryngeal. The tableau (5) below shows the interaction between these two constraints. Since input laryngeal features faithfully show up in the output forms, the constraint Ident-IO(Lar) must rank higher than No-Lar in normal phonology:

(5) I-O feature faithfulness⁴

/p'ul /	Ident-IO (Lar)	No-Lar
<u>p'ul</u>		*
<u>pul</u>	*!	

According to the tableau, being faithful to the input features is more important than avoiding marked features in regular phonology of Korean. However, the exact opposite situation is observed in the reduplicated CV forms.

3. Reduplicative domain

3.1. Infixal CV reduplication

In Korean, optional, partial reduplication of a light syllable (CV) adds a repetitive or durative meaning to the (usually) mimetic base. Partially reduplicated mimetic words either denote several short consecutive movements or sounds in one event, or extend the state or sound of the base. Data below are given with the dash "-" representing the morpheme boundary, and the dot "." syllable boundary. Reduplicants are underlined:

(6) Infixal CV reduplication

<i>base</i>	<i>gloss</i>	<i>reduplicated</i>	<i>gloss</i>
1. p ^h aŋ	'a bang'	p ^h a.- <u>pa</u> -ŋ	'two bangs in one event'
2. t'ok	'a drop'	t'o.- <u>to</u> -k	'two quick consecutive drops'
3. sak	'ducking'	sa.- <u>sa</u> -k	' <u>ducking with quick movements</u> '
4. k'o.lik '(stomach) growling'		k'o.li.- <u>li</u> -k	'growl two short instances at once'
5. hu.tak	'hurriedly'	hu.ta.- <u>ta</u> -k	'hurriedly with several loud footsteps'
6. p'u.cik '(tree branch) crack'		p'u.ci.- <u>ci</u> -k	'several cracking sounds in one event'

As shown underlined above, the shape of the RED is CV. The RED is attached to the rightmost syllable of the base to the exclusion of its final consonant, giving the process a look of infixation. The empty CV template fills its segmental content by copying the final syllable of the base minus the final consonant of the base. (CV cannot have been prefixed, in which case a base like *k'o.lik* should reduplicate as **k'o-ko.lik*, which is ill-formed.) In all reduplicated forms, the laryngeal feature from the base is lost only in the underlined reduplicant (that is, if there was any in the final syllable of the base to begin with)⁵. The input laryngeal

feature shows up intact in the output forms of the base as high ranking Ident-IO(Lar) requires (e.g. *p^ha.-pa-ŋ*); the relevant feature, however, is absent in the reduplicant (e.g. *p^ha.-pa-ŋ*).

3.2. Analysis - laryngeal feature loss in CV is emergence of the unmarked

The loss of laryngeal feature in the reduplicant can only be explained if marked features are banned in the reduplicative domain (McCarthy and Prince 1994b). To put it in terms of ranked constraints, No-Lar the marked-feature-banning constraint is ranked higher than the base-reduplicant-identity-governing constraint. The following tableau makes the point clear:

(7) Emerging effects of constraint No-Lar in CV RED

<u>/p^haŋ-RED/</u>	<u>Ident-IO(Lar)</u>	<u>No-Lar</u>	<u>Ident-BR(Lar)</u>
<u>p^ha-p^ha-ŋ</u>		**!	
<u>*p^ha-pa-ŋ</u>		*	*
<u>pa-p^ha-ŋ</u>	*!	*	*
<u>pa-pa-ŋ</u>	*!		

In the tableau, the first candidate *p^ha-p^ha-ŋ*, which shows faithful feature correspondence between the base and the reduplicant and thus gets no mark on the Ident-BR(Lar)⁶ constraint, fatally violates the constraint No-Lar and loses out. In the case of the fourth candidate *pa-pa-ŋ*, although it incurs neither No-Lar nor Ident-BR(Lar) violation, its unfaithfulness to the input features turns out to be fatal. With the third candidate faring worst then, the crown for the optimal survivor goes to the candidate *p^ha-pa-ŋ*.

From the tableau (5) in section 2, it was shown that laryngeal features are allowed in the output forms in normal phonology (i.e., in the onset position), and from the tableau (7) above, it is obvious that they are banned in the reduplicative domain. Although the ranking of the two constraints Ident-IO(Lar) and No-Lar remains unchanged in both domains, the markedness constraint No-Lar comes alive in the domain of reduplication, that is, No-Lar intervenes between the two, sacrificing the featureal identity between the base and the reduplicant. As a result, the markedness effect emerges. (McCarthy and Prince 1994a, 1994b, 1995).

4. Puzzle: some RED's do copy Lar features

My analysis, however, faces a little conflict when RED of a different shape in Korean, namely CVC, is considered. Observe the characteristics of the CVC RED below with specific attention to laryngeal feature copying, which present a striking contrast with the CV RED.

The CVC RED may be prefixal (a) or suffixal (b). When optionally prefixed, the base meaning is either emphasized or extended (e.g. *ki.ce* 'yesterday' -> *ki-ki.ce* 'the day before yesterday'). When suffixed, the CVC RED usually describes the quality or state of the subject of the sentence (usually occurs in the form of *adj.-hata* as in *al.t'al-t'al-ha.ta*):

(8) CVC reduplication

(a) prefixal reduplication

<i>base</i>	<i>reduplicated</i>	<i>gloss</i>
1. t'e.kul	t'ek.-t'e.kul	'rumble'
2. ti.kil	tiik.-ti.kil	'crowded'
3. t'hə.lim	t'həl.-t'hə.lim	'sour'
4. t'a.kul	t'ak.-t'a.kul(i)	'woodpecker'
5. ki.ce	kič.-ki.ce	'two days ago'

(b) suffixal reduplication

<i>base</i>	<i>reduplicated</i>	<i>gloss</i>
1. *al.t'al	al.t'al-t'al	'buzzed'
2. *a.sam	a.sam.-sam	'dazzled'
3. *twi.suŋ	twi.suŋ.-suŋ	'perturbed'
4. *mu.təm	mu.təm.-təm	'indifferent'
5. *o.coŋ	o.coŋ.-coŋ	'small'

Obvious from the data is the template shape (CVC) and, unlike the CV reduplication, the faithful laryngeal feature copying of the base onto the reduplicant in both cases (prefixal: *t'e.kul* --> *t'ek-t'e.kul*; suffixal: *al.t'al* --> *al.t'al-t'al*).

A tableau that reflects the behavior of the laryngeal features of the CVC reduplication has to now re-rank the constraints No-Lar and Ident-BR(Lar), yielding an opposit constraint ranking to that of the CV reduplication. (9) is such a tableau (As their behaviors are identical with respect to feature shedding in prefixal and suffixal CVC reduplication, I will discuss only the suffixal CVC below):

(9) Reverse ranking of constraints in CVC RED

/al.t'al-RED/	Ident-IO (Lar)	Ident-BR (Lar)	No-Lar
*al.t'al-t'al			**
al.t'al-tal		*!	*
al.tal-t'al	*!	*	*
al.tal-tal	*!		

Since the input features still have to have corresponding features in the output, the constraint Ident-IO(Lar) is left undominated, which forces the third and the fourth candidates that contain no laryngeal feature in the output form of the base to lose out, just as in CV reduplication. Because the real output is *al.t'al-t'al* (marked optimal with the index finger), the base-reduplicant identity-governing constraint Ident-BR(Lar) must now rank higher than No-Lar, incidentally eliminating the second candidate *al.t'al-tal* also. One simple explanation for this ranking conflict would be to say that constraint ranking is dependent on each reduplicative morpheme and that the behavior of the laryngeal feature will be dictated by whatever is required of the individual RED morpheme in the lexicon (Alderete et al 1996, McCarthy and Prince 1994b).

Although this conjecture correctly describes the data, it does not seem to explain why there is a split in feature copying over different reduplicants' size. Why feature loss in CV but retention in CVC reduplicant⁷? If the effects of the markedness constraint emerge uniformly everywhere in the reduplicative domain per se, one would expect to see the same behavior of laryngeal features with respect to copying in both the CVC and CV reduplicants. It may be possible to encode the information about the laryngeal feature copying for each morpheme. But this morpheme-by-morpheme feature copying runs into paradox: reduplicative morphemes normally lack their melody content, but each Korean CV and CVC reduplicative morpheme has to have the information whether the melody to be copied will include or exclude laryngeal features. That is, a very specific feature copying information has to be pre-specified for the general CV or CVC RED whether their future base they copy melody from turns out to have (or not have) laryngeal features. Without this kind of complication, the following section answers the question why there is a puzzling split behavior of the laryngeal features in Korean.

5. Proposal: derived "templates" and according RED behavior

I propose to solve this puzzle by adopting McCarthy and Prince's proposal that templates are derivable from the stem and the affix in regular phonology. Their specific claim is that affixes tend to be universally less marked while stems tolerate marked structure both in terms of syllable and feature structure. If a RED, some

morpheme that lacks its segmental melody, is classified as a type of affix in the lexicon, that specific RED will demonstrate its morphological affiliation by allowing no marked features. If a RED is classified as stem, on the other hand, it will tolerate I argue, or even force marked base features to appear in the reduplicant. According to this proposal, the infixal CV RED in Korean will have the classification of affix and the CVC RED of stem in the lexicon. The CVC RED then becomes a kind of stem-stem compounding with one of the member of the compound lacking its segmental melody. Below I will show the respective unmarked and marked status of affix and stem in normal phonology of Korean, relying on their size and feature requirement.

5.1 Evidence from normal phonology for RED=affix

The so called 'tV infixation' in regular phonology in Korean offers supporting evidence for the infixal RED taking on the characteristics of affix. In (10) below, data from tV infixation, (which invariably requires a heavy disyllabic mimetic base) is presented to be compared with the CV reduplication:

(10) Comparison of tV infixation and CV RED

<tV infixation>			<Infixal CV reduplication>		
<i>base</i>	<i>affixed</i>	<i>gloss</i>	<i>base</i>	<i>reduplicated</i>	<i>gloss</i>
1. t ^h əl.s'ək	t ^h əl.s'ə.- <u>l</u> ə-k	'flop'	1. t ^h aŋ	t ^h a.- <u>l</u> a-ŋ	'bang'
2. k'ol.k'ak	k'ol.k'a.- <u>l</u> a-k	'gulp/gluck'	2. t'ok	t'o.- <u>l</u> o-k	'(small object) drop'
3. hul.ləŋ	hul.lə.- <u>l</u> ə-ŋ	'over'	3. p'a.tik	p'a.ti.- <u>l</u> i-k	'teeth-grinding'
4. c ^h əl.ləŋ	c ^h əl.lə.- <u>l</u> ə-ŋ	'sudden drop without touching the ground'	4. cu.luk	cu.lu.- <u>l</u> u-k	'dribble'
5. t'al.kak	t'al.ka.- <u>l</u> a-k	'click'	5. pu.liŋ	pu.li.- <u>l</u> i-ŋ	'vrrrum'

As obvious from the compared data, the affix tV recaptures the characteristics of the CV reduplicant, by attaching to the stem before its final consonant. Both segments 't' and 'V' in this affix show an extreme case of unmarkedness, the 't' being so-called the unmarked coronal consonant in Korean (You-Cho), and the vowel 'V' harmonizing with the vowels of the stem. This is an exact example for McCarthy and Prince's claim that affixes are unmarked with respect to the feature content. The shape of the affix is also notable, being codaless, just like the CV reduplicative template. An obvious generalization then comes into the picture: the affix tV,

showing up in the normal phonology, and the CV of the reduplicative domain, are both be specified as affix in the lexicon. If these two are morphologically one, namely affix, then there is no surprise in their sharing phonological characteristics. Consequently, since we know that affixes are unmarked (or less marked) in Korean, eliminating the CV template and other ancillary feature specifications (i.e., banning specific features) for reduplicative morphemes in the lexicon may be possible⁸.

5.2 Evidence from normal phonology for RED=stem

In Korean, words have to be of a certain size in the output form. Functional words may have the shape of (C)V. Demonstratives and the mood marker *ta*, for example, may occur in the form of a light syllable:

<function words>

demonstratives	i 'this',	ki 'that',	cə 'that over there'
mood marker	ta:	ka -ta 'go,	cuk -ta 'die'

On the other hand, small size lexical words are lengthened in isolation, but heavy syllable words surface as they are with no lengthening or augmentation in the output form:

<lexical words>

underlying	CV	/k ^h o/ -> [k ^h o:] 'nose',	/i/ -> [i:] 'tooth/teeth'
underlying	CVC	/kat/ -> [kat] 'river',	/sot ^h / -> [sot] '(big) pot'

This observation brings in the notion of prosodic word, namely prosodic words have to satisfy a minimum size requirement, which is a heavy monosyllable in Korean. This claim makes an implicit prediction that prosodic words have to be at least a heavy monosyllable, which is borne out by some CVC prefixes in Korean that pattern with the prosodic words not with the suffixes (Kang 1991c). McCarthy and Prince argue for why specifying the size of the prosodic word is unnecessary. Using alignment constraints, they claim that the size of prosodic word is easily derivable from that of the stem. I will not go into the issue here, but interested readers are referred to their work (1994a).

5.3. Conclusion

Resorting to McCarthy and Prince's (1994a) claim that Stem=Pwd in size, where a stem may bear marked features more freely than affixes, I conclude that the Korean CVC RED is in fact classified as stem in the lexicon. If stems in Korean -- which is of a certain required size-- may bear marked features and if the CVC RED is a stem, the CVC RED is naturally allowed to bear marked features. Or, the CVC RED as a stem has to bear marked features present in the base. This is exactly what we have seen in the CVC prefixal and suffixal reduplication, resulting in higher ranking of Ident-BR than No-Lar. It is also shown that the unmarked nature of the CV RED with respect to features can be ascribed to the unmarked phonological characteristics of affixes. Since CV affixes are less marked and the CV RED demonstrates such characteristics, there may not be a reason to state the template of this particular RED in question as CV in the grammar. Simply the grammar needs recognize that it is a species of affix. Thus RED=affix specification in the lexicon will give us all the necessary results, namely the RED in question will apparently look like CV, and there will be no laryngeal feature copying.

6.1. A potential problem

Template elimination is certainly desirable when simplicity and economy of grammatical statement is considered. My proposal that the CVC RED is a stem, however, poses a potential problem. In a language with more than one stem-stem compounding reduplication, the simple dichotomy between RED=affix and RED=stem would not be sufficient to characterize all reduplicant types. In the very language of our discussion, Korean, a mimetic base has several options to entertain. It can be partially reduplicated by repeating the last syllable minus its coda as in CV reduplication or by repeating the whole first or last heavy syllable as in CVC reduplication. It can also repeat itself wholly, as in total reduplication (e.g. *t'e.kul-t'e.kul* 'rumble (repeated event)'). Since the CV and the CVC reduplication would exhaust all the available lexicon classification, total reduplication is left in limbo. It is also a stem-stem compounding reduplication, but how could the grammar recognize it? To remedy the situation, I propose two types of stem RED in the lexicon, namely, Stem_{min} and Stem_{max}. The default assumption would be that the Stem_{max} will dictate total reduplication and that Stem_{min} will be associated with the reduplication involving the specific minimal

stem size of each individual language. Note that I am not directly addressing the exact shape of RED. As the number of syllables in the base will vary, there can be no one template shape statable in the case of total reduplication any way. The following shows how Korean works according to this schema:

RED₁ (partial) - stem+affix: *p'a-pa-ŋ* (RED=affix)

RED₂ (partial) - minimal stem+stem compounding: *t'ek-t'e.kul* (RED=stem_{min})

RED₂ (total) - maximal stem+stem compounding: *t'e.kul-t'e.kul* (RED=stem_{max})
(or, *t'e.kul-t'e.kul*)

Although the exact nature of the operation of Stem_{max} and Stem_{min} goes beyond the scope of this paper, one conjecture is possible. If it is true that languages do not allow more than two kinds of stem+stem compounding reduplication (one being total and the other repeating whatever the minimal stem size of the individual languages), the lexicon specification for RED does not have to become any more complicated by allowing RED=stem₁, RED=stem₂, ... etc. I know of no language that allows such a pattern. Also, if RED=affix, RED=Stem_{max}, and RED=Stem_{min} are in fact exhaustive reduplicative entries in the lexicon, any more types of reduplication will have to resort to some other parameters than these classifications. One very convincing piece of evidence is attested in Korean, in which following parameters are attested to give varieties to reduplicative types (SJ Kim in preparation):

a) RED=affix	b) RED=Stem _{min}	c) Stem _{max}
edge-differing: <i>ki-kil.pʰi</i> (prefix) 'three days later'	<i>t'ek-t'e.kul</i> (prefix) 'rumble'	
<i>a.cu-cu</i> (suffix) 'condescending exclamation'	<i>o.coŋ-coŋ</i> (suffix) 'roundish (facial trait)'	
epenthetic red: <i>sa.l-i.li</i> 'softly melting'	feature-adding red: <i>u.taŋ-tʰaŋ</i> 'thumping'	
geminating red: <i>om.ma-mma</i>	onset-dropping red: <i>o.son-to.son</i>	
	syll-changing red: <i>s'ip.t'uk-kəp.t'uk</i>	
	vowel-changing red: <i>siŋ.kil-seŋ.kil</i>	
	syll-alternating red: <i>sa-sa.kən-kən</i>	

Amazingly, the different types of reduplication in Korean above all fall into the three reduplicative classifications.

6.2. Theoretical implications

The Korean case provides a support for McCarthy and Prince's proposal of deriving templates from the affix and the stem has the advantage of writing a simpler grammar. A grammar that works with fewer gadgets is a better grammar as long as it is equally capable of explaining facts. By eliminating templates, we also have fewer potential constraints, consequently simplifying the grammar. As some have shown (Urbanczyk's constraint $Afx \leq \partial$), the template size can be directly translated into a constraint. If the size effect is derivable, however, there is no reason to add more constraints to the grammar. I have modified McCarthy and Prince's claim about the emerging effect of markedness constraints in the reduplicative domain in such a way that the effects of emerging markedness constraints are not so uniformed as expected in the reduplicative domain per se; rather, the visibility of unmarkedness constraints is a direct result of the unmarkedness of affixes. This modestly predicts that if an unmarkedness effect is shown in the reduplicative domain in a language, it is likely that the RED examined is classified as RED=affix in the lexicon or behave like an affix in that language. Finally, my proposal provide support for the constraint No-Lar.

* I would like to thank professor Sharon Hargus and the WECOL audience for their helpful comments.

¹As in Cho (1994), I take the contrast to be three way, rejecting the underlying geminate analysis (cf. Jun 1993, Han 1996).

²Laryngeal features are lost syllable-finally, but this is not a relevant issue, since in the reduplicative domain, no marked features are allowed in the coda position anyway.

³If binary specification is adopted, the intuition that the plain obstruent is unmarked is lost. For the case of plain, binary specification requires a change from [-SG, -CG] to [+voice], whereas privative specification from nothing to [(+voice)]. The intervocalic change from nothing to [(+voice)] in the laryngeal node is an apparent case of voicing spread (e.g., assimilation), but from [-SG, -CG] to [+voice] is totally haphazardous. Thus I conclude that laryngeal features (in Korean at least) are better specified privatively.

⁴Again, the discussion concerns only the laryngeal features in the onset position. Although obstruents in coda position get neutralized in place and manner, they are irrelevant since all base codas are underlyingly plain stops in the reduplicative domain (originally observed by Lee 1992).

⁵This laryngeal feature loss appears only in monosyllabic bases, since in no disyllabic base does the onset of the second syllable, the melody of which is copied, is laryngeally marked. This itself is interesting, but I will ignore the issue here.

⁶To be exact, features are evaluated in terms of faithfulness even in the correspondence framework. As in the containment theory, the features in the reduplicant are evaluated with respect to its 'feature faithfulness' to the base in the correspondence theory. Reduplicant features do not have "correspondent" features in the base; rather, segments are evaluated for their feature faithfulness. Whether such constraints as Max-F and Dep-F and the like should evaluate individual features comprising a whole segment should bear important theoretical consequences, but the problem does not directly concern the issue under discussion.

⁷Reduplicants show a parallel behavior with respect to laryngeal feature copying in total reduplication in Korean. For more, see section 5.

⁸To some, tV infixation is a species of reduplication. If this claim turns out to be true, my proposal is weakened. I could still keep the proposal, however, by proposing that the unmarkedness is a nature of "affixes" universally without trying to derive it from the affixes in normal phonology of a specific language. Many more languages will need to be examined to decide this matter.

Bibliography

- Aldertre et al. 1996. Reduplication and segmental unmarkedness. Ms. University of Massachusetts, Amherst.
- Clements, Nicholas. 1985. The geometry of phonological features: Phonology II.
- Han, JeongIm. 1996. The phonetics and phonology of tense and plain consonants in Korean. Doctoral dissertation, Cornell University.
- Iverson and Kim. 1987. Underspecification and hierarchical feature representation in Korean consonantal phonology. CLS 23.2: 182-98.

-
- Jun, Jongho. 1993. Prosodic approach on Korean partial reduplication. In P. Clancy (ed.) *Japanese/Korean Linguistics*. Vol. 2. Stanford: CSLI. 130-146.
- Jun, Jongho. 1994. Metrical weight consistency in Korean partial reduplication. *Phonology* 11. 69-88.
- Kang, Ongmi. 1991. Word-internal prosodic words in Korean. *NELS* 22:243-257.
- Kim, J. Soohye (to appear). Emergence of the unmarked: Korean partial reduplication. *Proceedings of Chicago Linguistics Society* 32.
- Kim-Renaud, YoungKi. 1974. *Korean Consonantal Phonology*. Doctoral Dissertation, University of Hawaii.
- Lee, Jinseong. 1992. *Phonology and sound symbolism of Korean ideophones*. Doctoral dissertation, Indiana University.
- Lee, Youngsang. 1994. Onset analysis of Korean on-glides. In *Theoretical Issues in Korean Linguistics*. CSLI.
- Lombardi, Linda. 1991. *Laryngeal features and privativity*, ms., University of Massachusetts, Amherst.
- Lombardi, Linda. 1995. Laryngeal neutralization and alignment. In *Papers in optimality theory*. University of Massachusetts occasional papers 18. 250-384.
- Maddison, Ian. 1984. *Patterns of Sounds*. Cambridge University Press.
- McCarthy, John and Alan Prince. 1986. *Prosodic morphology*. Unpublished Ms, University of Massachusetts and Brandeis University.
- McCarthy and Prince. 1990. Foot and Word in Prosodic Morphology: The Arabic Broken Plural. In *Natural languages and Linguistics theory* 8. 209-283.
- McCarthy and Prince. 1993. Generalized Alignment. In Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology*. The Netherlands: Kluwer Academic Publishers. 79-153.
- McCarthy and Prince. 1994a. An Overview of Prosodic Morphology. Pt. 1&2: Template Form in Reduplication; Template Satisfaction. Talks at Utrecht Prosodic Morphology Workshop.
- McCarthy and Prince. 1994b. Emergence of the Unmarked: optimality in Prosodic Morphology. ed. by Merce Conzales. *Proceedings of the Northeast*

-
- Linguistic Society 24. Amherst, MA: Graduate Linguistic Student Association. Pp 333-379.
- McCarthy and Prince (1995). Faithfulness and Reduplicative Identity. In Papers in Optimality Theory. University of Massachusetts Occasional Papers 18. 250-384.
- Urbanczyk (to appear). Morphological templates in reduplication. NELS 26.
- You-Cho, Youngme. 1998. Syllable Structure Conditions in Korean and the OCP. International Conference of Korean Linguistics. 2. Ed. E. Baek.

Two Types of Bare Measure Phrases*

Murat Kural

UCLA

1. Introduction

In this paper, the term ‘Bare Measure Phrase’ (BMP) refers to a phrase of quantity that scales the event or the state that is identified by the verb. Examples for BMPs that are relevant for the discussion are given bold-faced in (1) and (2).

- (1) The ship sank **400 meters**
- (2) The athlete ran **400 meters**

BMPs have two crucial properties. First, they are not oblique phrases, i.e., they are not PPs or oblique Case-marked DPs. Thus, the measure phrases in (3) do not qualify as BMPs.

- (3) a. The ship sank **for 400 meters**
- b. The athlete ran **for 400 meters**

The syntactic status of oblique and PP measure phrases is very different from the status of DP measure phrases, especially in the ways they are licensed in the structure and the range of grammatical functions they can assume.

The second property of the BMPs is that they are not temporal phrases that denote duration, which is why the examples in (4) do not count as BMPs.

- (4) a. The ship sank **5 minutes**
- b. The athlete ran **5 minutes**

Intuitively, temporal duration is not an integral component of an event. Rather, it appears as an incidental and arbitrary correlate of the fact that events take place in real time. For example, whether one walks 400 meters within three minutes or three hours depends on one’s pace and/or the length of the breaks one might take along the way. The amount of walking remains the same either way, which is the 400-meter distance. Therefore, it would be reasonable to say that the extent of the walking event is measured by the distance that is traversed, instead of the time that it takes to cover the distance, and that the time that elapses is not the defining property of events like walking.

Note that there are two conditions that must be met for a verb to allow a BMP: (a) It must identify a scalable act (event or state), which excludes most achievement verbs such as *break* and *explode*, and (b) the scalable act identified by the verb must have a well-defined (though perhaps not formalized) unit of measurement, which excludes verbs like *burn*, *change*, *degrade*, *evolve*, *freeze*, *improve*, and *wrinkle*. It is also important that the verb not directly incorporate the

measurement itself, as is the case with verbs like *double* and *triple*.

Despite their surface similarities, the BMPs given in (1) and (2) display very different syntactic properties in terms of their capacity to passivize and to have independent reference. It will be argued in this paper that these differences are determined by the semantics of the verb in question. With Change of State (COS) verbs, as in (1), the BMP is a predicative constituent that forms a complex predicate with the verb, but with Change of Location (COL) verbs, as in (2), the BMP is a nonthematic argument of the verb.¹

2. Verb Classes

This section establishes Change of State (COS) verbs and Change of Location (COL) verbs as two independent verb classes that are distinct from unaccusative and unergative verbs.

2.1. *Change of State Verbs*

COS verbs are verbs that indicate transition to some state, such as *advance*, *grow*, *open*, *sink*, *stretch*, *shrink*, *tilt*, and *warm*. They share certain syntactic properties with unaccusative verbs like *appear*, *arrive*, *emerge*, *exist*, and *occur*, which describe the act or state of being present. For example, both verb types select *essere* 'be' auxiliary in the perfective tense in Italian, and allow *ne*-cliticization (Burzio 1986). However, COS verbs do not belong in the same class as unaccusatives. Perhaps the most obvious difference between the two verb types is that COS verbs can freely transitivity in English, as in (5), while unaccusative verbs cannot, as in (6).

- (5) a. The enemy sank the ship
 b. Bill warmed the milk
 c. Sue shrank her sweater
- (6) a. *The magician appeared the rabbit
 cf. The magician made the rabbit appear
 b. *Bill occurred an accident
 cf. Bill made an accident occur
 c. *God existed the universe
 cf. God made the universe exist

Second, COS verbs do not allow *there*-insertion, as in (7), whereas unaccusative verbs do, as in (8).

- (7) a. *There sank three ships (in the harbor)
 b. *There warmed some milk (in the stove)
 c. *There shrank a few sweaters (in the dryer)
- (8) a. There appeared three rabbits (on the stage)
 b. There occurred many accidents (on the freeway)

- c. There emerged a few problems (in the project)

Third and most relevant for the discussion at hand is the fact that COS verbs allow BMPs, but unaccusative verbs do not, as seen in (9) and (10), respectively.

- (9) a. The ship sank 400 meters
 b. The milk warmed 20 degrees
 c. The sweater shrank two sizes
 (10) a. *The toothpaste appeared five inches
 b. *The novel emerged 250 pages
 c. *Rainfall occurred three inches

The sentences in (10) describe situations where the use of the BMP is essentially plausible, yet the unaccusative verbs fail to license BMPs. They are not acceptable in cases where the toothpaste that is squirted out of the tube is five inches long, the novel that emerges after the writing is 250 pages, and three inches of rain occurs during the storm. Since each situation is pragmatically well-formed, there must be structural reasons that unaccusative verbs cannot take BMPs. This suggests a difference between COS verbs and unaccusative verbs in terms of the way their VP architecture is organized. In the traditional VP structure of unaccusative verbs proposed by Burzio (1986), the complement position is reserved for the thematic argument (or its trace), which means that the complement position is not free to host any BMPs. As a result, VPs headed by unaccusative verbs do not allow BMPs, as suggested by (10).²

2.2. *Change of Location Verbs*

COL verbs indicate motion to or from a location, such as *drive*, *fly*, *jump*, *run*, *swim*, and *walk*. Although not crucial to the main point of this paper, it must be pointed out that COL verbs form a distinct class that excludes unergative verbs such as *cry*, *dream*, *laugh*, *speak*, and *think*, which very often describe acts of production. The two verb types syntactically behave the same way in many respects, such as the selection of the *avere* 'have' auxiliary in the perfect tense in Italian and not allowing *ne*-cliticization. However, there is a significant distinction between them: COL verbs typically allow transitivity, but unergative verbs do not, as shown in (11) and (12), respectively.

- (11) a. Bill ran the horses across the field
 b. Sue walked me to the car
 c. John flew the plane over the Atlantic
 (12) a. *The clown laughed the children
 cf. The clown made the children laugh
 b. *The movie cried the audience
 cf. The movie made the audience cry

- c. *The professor thought her students
- cf. The professor made her students think

This particular asymmetry suggests that COL verbs do not belong in the same class as unergative verbs.

2.3. *The Nonthematic Complement Position*

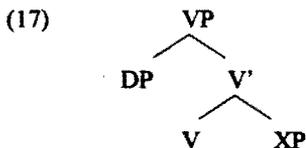
As two typologically distinct verb classes, COS and COL verbs have a number of properties in common, even though they differ in crucial ways. One striking similarity between them is that they are both capable of taking BMPs.

- (13) a. The ship sank 400 meters
- b. Bill grew two inches
- c. The picture tilted 30 degrees
- (14) a. The athlete ran 400 meters
- b. Sue walked a few miles
- c. The cat jumped ten feet

However, they differ with respect to their ability to license cognate objects: COS verbs do not allow them, as in (15), but COL verbs do, as in (16).

- (15) a. *The ship sank a great sink
- b. *The milk warmed a quick warmth
- c. *The sweater stretched a long stretch
- (16) a. The athlete ran a great run
- b. Sue walked a long walk
- c. The cat jumped a quick jump

The VP architecture assumed in this work for COS and COL verbs ascribes a similar internal structure, schematically represented in (17) below.



The ability of a verb to take a BMP indicates that its complement position, XP, is available for constituents that do not thematically relate to the verb, which is also the case with cognate objects and resultative phrases. It will be argued in section 5 that the difference between COS and COL verbs in terms of their capacity to take cognate objects, as seen in (15) and (16), is due to the syntactic nature of the type of complement these verbs can take.

3. Two Types of Bare Measure Phrases

Although they appear to be quite similar at the surface, the BMPs licensed by COS and COL verbs display two key differences in the syntax:

- (a) COS verbs do not allow BMPs to passivize, COL verbs do.
- (b) COS verbs do not allow specific BMPs, COL verbs do.

Each property is illustrated and discussed below.

3.1. *The Passivizability of BMPs*

BMPs cannot become the subjects of passives with COS verbs, but they do with COL verbs, which is shown in (18) and (19), respectively.

- (18) a. *400 meters were sunk by the ship
- b. *Two inches were grown by Bill
- c. *30 degrees were tilted by the picture
- (19) a. 400 meters were run by the athlete
- b. A few miles were walked by Sue
- c. Ten feet were jumped by the cat

There are two reasons that the inability of BMPs to passivize with COS verbs cannot simply be due to the inability of unaccusative verbs to passivize in most languages. First, as pointed out in section 2.1, COS verbs are not in the same class as unaccusative verbs. They behave differently with respect to transitivization, *there*-insertion, and the licensing of BMPs. Second, BMPs do not passivize with COS verbs in Turkish either, as shown in (20), even though Turkish allows COS verbs to form impersonal passives when there is no BMP argument, as in (21).³

- (20) BMP subject passives:
 - a. *50 metre rıhtımda (gemi tarafından) bat-ıl-dı
 meter harbor-LOC ship by sink-PASS-PAST-3SG
 ‘50 meters were sunk in the harbor (by the ship)’
 - b. *[iki yaş] [bu yetimhanede] (Ahmet tarafından) büyü-n-dü
 two age this orphanage-LOC A. by grow-PASS-PAST-3SG
 ‘Two years are grown in this orphanage (by Ahmet)’
- (21) Impersonal (generic) passives:⁴
 - a. Rıhtımda 50 metre (*[gemi tarafından]) bat-ıl-ır
 harbor-LOC meter ship by sink-PASS-AOR-3SG
 ‘There is sunk 50 meters in the harbor (by the ship)’
 - b. [Bu yetimhanede] [iki yaş] (*[Ahmet tarafından]) büyü-n-ür
 this orphanage-LOC two age A. by grow-PASS-AOR-3SG
 ‘There is grown two years in this orphanage (by Ahmet)’

Since COS verbs can in principle be passivized in Turkish, the failure of the BMPs to become derived subjects in the English (18) and Turkish (20) cannot be due to the way COS verbs react to passivization. Rather, it must be the result of some specific property of the BMPs in these structures.

3.2. *The Specificity of BMPs*

In Turkish, COS verbs do not allow their BMPs to bear the accusative Case morphology, which is closely associated with the specificity of the direct object (Enç 1991). The strict prohibition against accusative marking suggests that these BMPs cannot be specific expressions in Turkish.

- (22) a. Gemi 50 metre(*yi) battı
 ship meter-(ACC) sank
 'The ship sank (the) 50 meters'
 b. Ahmet onbeş santim(*i) büyüdü
 A. fifteen centimeter-(ACC) grew
 'Ahmet grew (the) fifteen centimeters'

By contrast, BMPs that are licensed by COL verbs may bear the accusative Case morphology, indicating that they are capable of being specific.

- (23) a. Atlet 400 metre(yi) koştu
 athlete meter-(ACC) ran
 'The athlete ran (the) 400 meters'
 b. Ayşe iki metre(yi) atladı
 A. two meter-(ACC) jumped
 'Ayşe jumped (the) two meters'

Neither sentence in (22) is acceptable in contexts that are established in a way that would force the specific reading. For example, one can imagine a 'sinking competition' that is parallel to the running example in (23a), in which ships are being monitored for the depth that they sink. Such a context makes no difference in terms of the prohibition against accusative-marked BMPs in (22a). The BMPs still have to be nonspecific regardless of such contexts, and they cannot bear accusative morphology.

A similar restriction against the specificity of BMPs that are licensed by COS verbs also obtains (to an extent) in English with respect to pronominalization. As seen in (24) below, BMPs that are complements of COS verbs cannot be replaced by pronouns, while the BMP complements of COL verbs in (25) can.

- (24) a. *The ship sank it quickly (it = 400 meters)
 (Context: The sea is 400 meters deep here)

- b. *The milk very quickly warmed it in the microwave (it = 20 degrees)
(Context: We need the milk 20 degrees warmer)
 - c. *The sweater shrank it in the dryer (it = two sizes)
(Context: The sweater is two sizes smaller now)
- (25)
- a. The athlete ran it during the race (it = 400 meters)
(Context: The distance is 400 meters)
 - b. Sue walked it last Monday (it = 20 miles)
(Context: It is 20 miles to my house)
 - c. The cat jumped it to get away (it = five feet)
(Context: The wall is five feet high)

Pronouns like *it* cannot refer to nonspecific individuals, nor can they introduce nonspecific elements into the discourse. For example, the subject of the sentence *It runs over cats* cannot refer to some nonspecific bus that runs over cats. The use of *it* is felicitous only if the bus is a familiar one in the context or a generic term that refers to all buses in general. Even when pronouns have indefinite discourse antecedents, as in *I would like to buy a book. I will show it to you when I do*, what they refer to is still a specific item whose identity is not established yet. That is, *I will show it to you* means *I will show [the book that I buy] to you*. From this perspective, the inability of BMPs to pronominalize in (24) shows that the restriction against the specificity of BMPs with COS verbs is not a phenomenon unique to Turkish.

The generalization that has emerged so far in this paper is that a BMP can neither be a specific expression nor a derived subject when it is generated as the complement of a COS verb. The next section will argue that the failure of such BMPs to be derived subjects or specific expressions stems from their predicate status in these structures, which is ultimately determined by the semantics of the COS verbs. It will be suggested in section 5 that the type of syntactic relation between the verb and its complement that is imposed by the semantics of the verb also accounts for the inability of COS verbs to license cognate objects.

4. The Semantics of COS and COL verbs

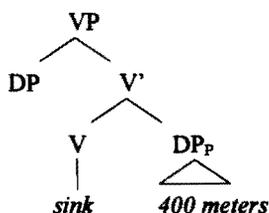
By their definition, COS verbs describe transition from one state to another, and as such, they designate a new state for the subject. For example, a ship that sinks becomes a *sunk ship*, or milk that warms becomes *warm milk*. The BMPs in *sink 400 meters* and *warm 20 degrees* indicate the extent of the transition to the state of being sunk and warm(er). In this sense, a ship that sinks 400 meters becomes a *400-metersunk ship* and milk that warms 20 degrees becomes *20-degree warm(er) milk*. Since they specify the state that the subject is in, BMPs licensed by COS verbs do not refer to any quantity that can exist independently of the subject. For example, there is no 400 meters that can stand apart from the sinking event in *sink 400 meters*, just as there is no 20 degrees that can exist outside the warming event in *warm 20 degrees*.

The semantics of COL verbs is very different from COS verbs. They do not indicate transition to a different state, so an athlete that runs does not become **a run athlete*. The BMP in *run 400 meters* refers only to the distance that is traversed by the subject, but not to any state that holds for the subject. Therefore, an athlete that runs 400 meters does not become **a 400-meter run athlete*, cf. **a run athlete*. With COL verbs, the subject maintains the same state throughout the event, and the BMP is merely the measurement of some quantity that exists independently of the subject or the event identified by the verb.

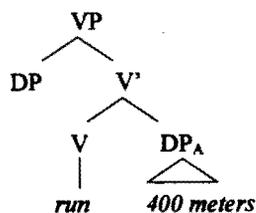
This particular semantic contrast shows that BMPs serve different grammatical functions with COS and COL verbs. Specifically, the BMPs of COS verbs act like secondary predicates in the way that they contribute to the state of the subject asserted by the main (COS) verb. This suggests that these BMPs are predicative constituents, which allows them to form a complex predicate with the verb, e.g., *The ship 400-meter-sank*. By contrast, the BMPs licensed by COL verbs are arguments of the verb, because the verb mediates a relationship between the subject and the BMP. In more concrete terms, the verb *run* establishes the relationship between *the athlete* and *400 meters* in *The athlete ran 400 meters*, which suggests that the BMP *400 meters* is an argument of *run*.

The main thesis of this paper is the following: The BMP that is licensed by a COS verb is a predicate DP, which forms a complex predicate with the verb to assert the state that the subject is in. On the other hand, the BMP that is licensed by a COL verb is an argument DP, which is selected as a nonthematic argument generated in the complement position. This distinction is expressed in (26) with the use of the subscripts 'P' for 'predicate' and 'A' for 'argument'.

(26) a. COS verbs:



b. COL verbs:



As argued by Stowell (1989), it is desirable to maintain a categorial distinction between nominal phrases that are predicates and arguments, and the labels DP_P and DP_A in (26) presumably correspond to a categorial distinction, albeit one that is left unspecified here.

The crucial point here is that since the predicative DP_P in (26a) is not an argument, it cannot become the subject of a passive form, just as the predicative DP complements of verbs like *become*, *elect*, and *name* do not:

- (27) a. Bill became a student
 b. *A student was become (by Bill)
- (28) a. They elected Mary the president
 b. *The president was elected (of) Mary
- (29) a. We named John the head of the committee
 b. *The head of the committee was named (of) John

On the other hand, specificity is not a notion that is applicable in the case of a predicate, so the predicative BMPs cannot be specific or nonspecific. As a result, they cannot bear the accusative morphology in Turkish, or be replaced by a pronoun in English. The argument DP_A in (26b), however, is generated as the complement of the verb, and although it bears no thematic role, it is an argument, and so it is capable of becoming the subject of a passive. For the same reason, it can also be specific or nonspecific, meaning that it can bear the overt accusative morphology in Turkish.

5. Conclusion and Implications

This paper has shown that the BMPs selected by COS verbs and COL verbs differ with respect to their capacity to be (a) derived subjects in the passive construction, and (b) (specific) direct objects in the active. It is argued that the difference is determined by the syntactic properties of the two BMP types, which in turn, is derived from the semantic properties of the verbs that license them. This is a significant conclusion because it allows children to determine the status of any given BMP without exposure to any overt evidence. They only need access to the information regarding whether the verb indicates a change of state or a change of location, which is clearly a key component of the meaning of the verb in question and hence, it is immediately accessible for children.

An important implication of the categorial distinction drawn between the two BMP types is that it makes it possible to account for the distribution of cognate objects with COS and COL verbs. One can plausibly argue that COS verbs cannot take arguments as complements because they predicate over their subjects and assert a new state. As a result, these verbs would be restricted only to predicate complements. Cognate objects may have complex internal structure, and they do not assert any state for the subject. As such, they appear to have are plausibly analyzed as arguments. That is, they are cases of DP_{AS} rather than DP_{PS} , and so COS verbs cannot license them as their complements.

- (15) a. *The ship sank a great sink
 b. *The milk warmed a quick warmth
 c. *The sweater stretched a long stretch

On the other hand, COL verbs allow arguments as their nonthematic complements, which is why they can take cognate objects.

- (16) a. The athlete ran a great run
 b. Sue walked a long walk
 c. The cat jumped a quick jump

Another implication of the predicate/argument distinction between the BMP types is that it is applicable in the case of another verb class, which comprises of verbs of measurement such as *weigh* and *cost*.

- (30) a. The box weighs 25 pounds
 b. The dress costs 1000 dollars

As has been noted in the literature, these verbs do not allow their BMPs to become derived subjects in passives.

- (31) a. *25 pounds is/are weighed by the box
 b. *1000 dollars is/are cost by the dress

The BMPs of measurement verbs are similar to the BMPs of COS verbs in terms of their predicative nature. They form a complex predicate with these verbs, and make an assertion about the state of the subject, e.g., *The box 25-pound-weighs* and *The dress 1000-dollar-costs*. Therefore, they also appear to be predicative constituents that are not capable of being derived subjects or specific phrases.

Notes

* I would like to thank Teun Hoekstra, Hilda Koopman, Dominique Sportiche, and all the participants of the Friday seminars at UCLA in the past two years, as well as Sandy Chung and other members of the audience at the WECOL 96 meeting at University of California, Santa Cruz for their helpful criticism and suggestions. A more complete and detailed discussion of the construction presented in this paper and its consequences on verb typology can be found in Kural (1996). An earlier version of this paper was presented at ESCOL '96 at UNB, Saint John.

¹ It has been argued by Stowell (1989) that the predicative nominal XPs are NPs, while the argument nominal XPs are DPs. This work will follow Stowell's insight that the predicative-argumental distinction has categorial basis, but it will do so without committing to any exact label for either category.

² The fact that COS and unaccusative verbs share certain characteristics, such as *essere* 'be' auxiliary selection and *ne*-cliticization in Italian (Burzio 1986), does not necessarily mean that they must be placed in the same class. For example, English uses the same *have* auxiliary for both unergative and unaccusative verbs, but this is normally not taken as evidence that all intransitive verbs belong in the same class in

English. Likewise, it is entirely possible that both verb classes fulfill the basic requirements of *essere* auxiliaries and *ne*-cliticization without having identical VPs.

³ The choice between *-il-* and *-n-* is phonologically determined.

⁴ Passives of unaccusatives and COS verbs must be impersonal and generic: They cannot refer to a specific event (Sezer 1991), and the understood subject must be human (Knecht 1985, Özkarağöz 1986). The sentences in (21) roughly translate as *One sinks 50 meters in this harbor* and *One grows two years in this orphanage*

References

- Burzio, L. 1986. *Italian Syntax*. Reidel Publications.
- Enç, M. 1991. The Semantics of Specificity. *Linguistic Inquiry* 22, 1-25.
- Knecht, L. 1985. *Subject and Object in Turkish*. Doctoral dissertation. MIT.
- Kural, M. 1996. *Verb Incorporation and Elementary Predicates*. Doctoral dissertation. UCLA.
- Özkarağöz, I. 1986. *The Relational Structure of Turkish Syntax*. Doctoral dissertation. University of California, San Diego.
- Sezer, E. 1991. *Issues in Turkish Syntax*. Doctoral dissertation. Harvard University.
- Stowell, T. 1989. Subjects, Specifiers, and X-bar Theory. In M. Baltin and T. Kroch, eds., *Alternative Conceptions of Phrase Structure*. University of Chicago Press.

Focus in the Future and the Thetic/Categorical Distinction*

Felicia A. Lee

University of California at Los Angeles

0. Overview

San Lucas Quiavini Zapotec (SLQZ), an Otomanguean language spoken in Oaxaca, Mexico, is a VSO language that allows the option of SVO word order. SVO word order, which is used to give contrastive focus to the subject, however, is disallowed in sentences with matrix verbs marked with the Definite aspect, which is used to strong speaker presupposition about future events. This paper proposes a correlation between the emphatic future readings of the Definite aspect and the ungrammaticality of preverbal subjects with this aspect: preverbal subjects raise to the specifier of a preverbal focus projection, and represent subjects of categorical judgments (in the sense of Kuroda 1972, elaborating upon ideas first proposed by Franz Brentano in 1874);¹ Definite verbs obligatorily raise to the head of the focus projection, thus representing thetic judgments. Because thetic judgments represent emphasis given to events and situations, rather than to individuals—and because SLQZ disallows more than one focused constituent per sentence—preverbal subjects are barred from raising to either focus or topic positions in sentences with Definite aspect.

SLQZ allows a choice between two aspects to express future events: the Irrealis and the Definite. Use of the Definite aspect (2) implies stronger speaker belief that an event will take place than use of the Irrealis (1):²

- | | |
|---|--|
| (1) I-to'oh Gyeihlly ca'rr
irr-sell Mike car
"Mike will sell the car" | (2) S-to'oh Gyeihlly ca'rr
def-sell Mike car
"Mike will (definitely) sell the car" |
|---|--|

While preverbal subjects are freely allowed in sentences whose matrix clause verbs are marked with Irrealis aspect (3), they are ungrammatical in sentences with matrix clause verbs with Definite aspect (4):

- | | |
|---|---|
| (3) Gyeihlly i-taz Lieeb
Mike irr-hit Felipe
"MIKE will hit Felipe" | (4) *Gyeihlly s'taz Lieeb
Mike def-hit Felipe
"MIKE will definitely hit Felipe" |
|---|---|

* Funding for this project was provided by a grant from UCLA's Institute of American Cultures/Chicano Studies Research Center, and by support from the UCLA Linguistics Department. SLQZ data was provided by Rodrigo Garcia. Thanks are due to Donka Farkas, Brenda Kennelly, Hilda Koopman, Pamela Munro, and the participants in UCLA's American Indian Seminar and Syntax/Semantic Seminar, and the audience of WECOL 1996 for their suggestions and questions. Any remaining errors are my own.

¹Brentano, Franz (1973) *Psychology from an Empirical Point of View*. Translated by Antos C. Rancurello, D.B. Terrell, and Linda L. McAlister from *Psychologie vom empirischen Standpunkt* (1874, 1924).

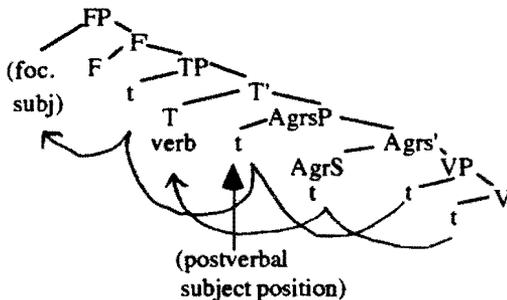
²SLQZ aspect markers often have more than allomorph: Irrealis aspect can be marked by *i-*, *ch-*, *l-*, or *gu-* prefixes; Definite aspect appears as either *z-* or *s-*. SLQZ, like other Zapotecan languages, overtly reflects aspect, rather than tense, in its verbal morphology (although certain aspect markers also encode tense features).

This paper proposes a correlation between the structural constraints and semantic features of the Definite aspect. This correlation is based on a difference in judgment type expressed by clauses with Definite verbs on one hand, and those with preverbal subjects on the other: Clauses with Definite verbs obligatorily represent *thetic judgments* (that is, speaker attention to the existence or denial of an eventuality), while clauses with preverbal subjects represent *categorical judgments* (that is, speaker recognition of an entity and that entity's relation to a predicate).

Here I should clarify the definition of 'judgment' assumed in this paper: following Kuroda (whose theories will be elaborated upon in detail in the following section), I use "judgment" to indicate a speaker's perception of what is salient in the description of an event. Thus, in categorical judgments, the speaker "judges" the individual involved to be more salient than the event itself; while the opposite holds true in *thetic judgments*. This concept is not related in any way to the concept of grammaticality judgment.

This difference in judgment type is realized in the syntax of SLQZ in the following way: Irrealis verbs (as well as those with aspects other than Definite) remain in TP, thus allowing subjects to raise to pre-TP/IP focus or topic positions, as seen in Figure 1:

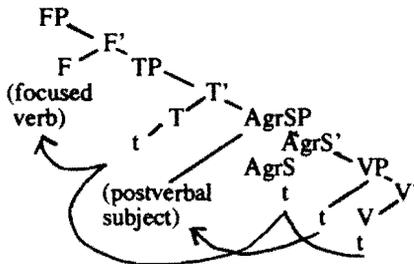
(Figure 1)



Focused (preverbal) subjects. (Irrelevant projections omitted)

On the other hand, Definite verbs, which represent assertion of the existence of an eventuality (focus of the predicate) raise to the head of the pre-IP Focus projection. Since SLQZ disallows more than one focused element per clause, subjects are blocked from raising to the specifier of the Focus position. (Figure 2):

(Figure 2)



Focused verb (=Definite aspect construction)

Sincethetic judgments are predicated on eventualities rather than individuals, subjects of these constructions are blocked from raising to either Focus or Topic for semantic reasons as well: by raising to either of these positions, the subject would receive extra semantic prominence in conflict with thethetic judgment type.

The correlation between the choice of focused constituent and the resulting judgment type expressed can be generalized as follows:

- Filled specifiers of FocusP represent subjects of categorical judgments
- Filled heads of FocusP result inthetic judgments

The body of the paper will be structured as follows: in Section 1, I will present a brief summary of past linguistic applications of thethetic/categorical judgment distinction. In Section 2, I will outline the behavior of syntactic focus (and some of its common uses) in SLQZ. In Section 3, I will further develop the correlation between argument focus and categorical judgments, and verbal focus andthetic judgments. Finally, in Section 4 I will use the constraints on interpretation of indefinites such as "someone" and "a person" in clauses with Definite verbs to support the proposal that Definite clauses obligatorily representthetic judgments.

1. Background

Kuroda (1972), updating ideas first proposed by Franz Brentano in 1875, proposes the distinction betweenthetic and categorical judgments as a means of articulating differences in perception of events that cannot be described in terms of differences in logical interpretation. He cites as an example the difference in speaker perspective expressed by an active sentence, such as *The Greeks defeated the Persians* versus those expressed by its passive counterpart, *The Persians were defeated by the Greeks*. While both sentences share the same argument structure, the former sentence (potentially) draws attention to the agent, while the latter draws attention to the theme.

Following Brentano, Kuroda divides such differences in perception into two basic types: CATEGORICAL judgments and THETIC judgments. Categorical judgments are constructed in two separate stages: they first require recognition of an entity that is to be the subject of the sentence, and second, affirmation or denial of what the predicate says about this subject. Thus, categorical judgments are also known as "double judgments".

In contrast, *thetic*, or "single," judgments are constructed in a single stage: "the recognition or rejection of material of a judgment," in Kuroda's terms. This notion will be made clearer in the examples below.

Kuroda proposes that the *thetic/categorical* distinction is reflected in the syntax of Japanese. He gives as an example the contrasting readings that result when a simple sentence such as *the dog is running* is expressed with the topic marker *wa* and the marker *ga*.

The *wa* example (5), Kuroda argues, represents a categorical judgment:

- (5) Inu *wa* hasitte iru
 "The dog is running" (categorical) (Kuroda 1972)

By using the topic marker *wa*, the speaker expresses the idea that the dog (which must be a specific dog already entered into the discourse) is the most salient part of the sentence.

It should be made clear that the notion of "subject" in this perceptual framework is independent of the standard definition of grammatical or thematic subjects: an object may be the subject of a categorical judgment, as in the following example (which Kuroda suggests can be best translated into English as a passive construction):

- (6) Neko wa inu ga oikakete iru
 "The cat is being chased by a dog"

In this case, the syntactic object *neko*, "cat" is marked with *wa* and is interpreted as the most salient participant in the event.

Thus, the term "subject" in the remainder of this paper will refer strictly to the salience of a particular argument within a categorical judgment, not to its syntactic or thematic role, unless explicitly noted.

In contrast to (5), the sense of salience of "dog" is lost if *ga* is used in place of *wa*, as seen in (7):

- (7) Inu *ga* hasitte iru
 "A/the dog is running" (*thetic*) (Kuroda 1972)

(7) would be used if someone saw a dog and said "a dog is running" (or "the dog/Fido is running") This *thetic* reading expresses the idea that the existence of an event that happens to involve a dog, not the dog itself, is particularly salient.

Note that while subjects of categorical judgments (such as the dog in (5)) must be presupposed entities, the same is not true for participants in eventualities expressed by *thetic* judgments: These may either be presupposed, specific entities or indefinites. This distinction will be a crucial diagnostic for *thetic* judgments later in this paper.

2. Syntactic Focus in SLQZ

Before moving to the direct correlation between focus and judgment type in SLQZ, I will provide some background on the behavior and uses of syntactic focus in SLQZ. Focused constituents in SLQZ appear in a preverbal focus position (indicated by boldfaced text in (9)), and preverbal subjects generally receive contrastive focus readings:³

³ There is ample evidence that in SLQZ, as in many other languages, *wh*-words occupy the specifier of the Focus projection. *Why*, however, seems to occupy a higher projection, since it is the only *wh*-word that allows other material to appear in preverbal focus position (as seen in (8)).

(8) **Zhini' b-ta'z Gyeihlly Lieeb?**
 why perf-hit Mike Felipe
 "Why did Mike hit Felipe?"

(9) **Zhini' Gyeihlly b-ta'z Lieeb?**
 why **Mike** perf-hit Felipe
 "Why did **MIKE** (rather than someone else) hit Felipe?"

I will assume (in the spirit of Rizzi 1995) that the focus features are checked when a focused constituent raises into FocusP, a functional projection above IP/TP.

Answers to argument wh-questions are also focused. The answers to both the subject and object argument wh-questions in (10) and (11) obligatorily appear in the preverbal focus projection:

(10)
 Q. **Tu b-dauhw comieed?**
 who perf-eat food?
 "Who ate the food?"

(11)
 Q. **Xi i-to'oh Gyeihlly?**
 what irr-sell Mike
 "What will Mike sell?"

A1. **Gyeihlly** (b-dauhw comieed)
Mike (perf-eat food)
 "Mike did"

A. **X:ca'rr Gyeihlly** i-to'oh Gyeihlly
poss-car Mike irr-sell Mike
 "Mike will sell his car"⁴

A2. **#B-dauhw Gyeihlly** comieed
 perf-eat **Mike** food
 "Mike ate the food."

While A1 (with a preverbal subject) is an appropriate answer to the question "Who ate the food?", A2, with a postverbal subject, while perfectly grammatical as an independent sentence of SLQZ, is not a felicitous answer to (10). The same holds true for object wh-questions such as (11): Here, as in (10), the answer to the question (in this case "Mike's car") must appear in preverbal position.

In contrast, subjects of event wh-questions must appear postverbally, as seen in (12):

(12)
 Q. **Xi b-e:e:i'ny Gyeihlly?**
 what perf-do Mike?
 "What did Mike do?"

A1. **B-zhu:u:u'nny Gyeihlly** loh Lieeb
 perf-run **Mike** from Felipe
 "Mike ditched Felipe"

A2. **#Gyeihlly b-zhu:u:u'nny** loh Lieeb
Mike perf-run from Lieeb
 "Mike ditched Felipe"

In questions such as (10) and (11), in which the subject or object in an event is being questioned, the subject or object must be fronted in the answer. In questions such as (12),

⁴ SLQZ, like other Zapotec languages, allows lexical nouns to be bound by other lexical nouns, but not by pronouns (Munro 1994, Black 1994)

in which a whole event, rather than a participant in the event, is being questioned, neither the subject nor object may be fronted; rather, the answer must assume canonical VSO form.

3. Correlations Between Syntactic Focus and Thetic/Categorical Judgments

The correlation between the *wh*-question type and the possibility of focused arguments brings to mind the thetic/categorical distinction outlined earlier. In this section, I will try to make some of these correlations explicit.

First, consider the case of argument *wh*-questions. When a speaker asks an argument *wh*-question, he or she makes two judgments: one, that an event occurred, and two, that he/she presupposes the existence of a specific participant in the event and wants more information about it. Thus, I will assume the following:

- Answers to argument *wh*-questions represent *categorical* judgments

Consistent with this, I will further assume that focused arguments in SLQZ correspond to subjects of categorical judgments.

On this point, my analysis diverges slightly from Kuroda's (or, more specifically, the behavior of arguments in *wh*-questions differs in SLQZ and Japanese): while he shares my belief that answers to argument *wh*-questions are focused, he notes that the syntactic subjects of such questions, if not *wh*-elements themselves, appear with the the topic marker *wa*, which as previously mentioned, marks them as subjects of categorical judgments:

- (13) (a) Fido wa nani o oikakete iru ka
 "What is Fido chasing?"
 (b) Fido wa *neko* o oikakete iru
 "Fido is chasing a cat"

Thus, if Kuroda's analysis is correct, argument *wh*-questions in Japanese, as in SLQZ, reflect categorical judgments, but differ from SLQZ in that the subject of the categorical judgment is always the syntactic subject, rather than the argument being questioned. Thus in Japanese, topics, but not focused constituents, can be construed as subjects of categorical judgments.

I will not attempt to propose an alternative analysis for the Japanese data. I will argue, however, that because of the syntactic and semantic prominence given to answers of argument *wh*-questions in SLQZ, that it seems reasonable that these represent subjects of categorical judgments. Furthermore, following a number of accounts on the semantics of *wh*-questions (Groenendijk and Stokhof 1983, Brennan 1995, Lee 1994), I assume that answers to normal (as opposed to rhetorical or surprise) *wh*-questions generally come from presupposed sets of entities established by some conversational context. Thus, argument *wh*-questions can't be said to be existential (=thetic) expressions introducing new entities into the discourse. Rather, the "new" information that answers to argument *wh*-questions provide is the exact choice from among that set that makes a proposition true.

In contrast, answers to event *wh*-questions such as (11) can only appear in standard VSO form; none of their arguments may be focused. In these cases, the speaker asking the question makes only one judgment: something happened, and he/she wants more information about it. I will thus assume the following:

- Answers to event *wh*-questions represent *thetic* judgments

Further evidence for a correlation between argument focus and categorical judgments comes from the interpretation of the SLQZ noun *bu:unny*, "person/one", as a preverbal and postverbal subject.

The noun *bu:unny* has a number of uses in SLQZ. In preverbal position, it is interpreted as a generic term for "people/person", as seen in (14). (Bare nouns in SLQZ can be interpreted as either singular or plural):

- (14) **Bu:unny** g-auw buhdy, burr g-auw gyiihzh
person irr-eat chicken donkey irr-eat grass
 "People/a person will eat chicken, donkeys will eat grass"

In postverbal position, on the other hand, it is interpreted as an impersonal subject (15):

- (15) R-u:ally **bu:unny** liebr ira'ta zhih tyenn g-ahcbe:e'-ru' bu:unny
 hab-read **person** book every day because irr-learn-more person
 "One reads books every day in order to learn more"

My consultant also uses sentences with postverbal *bu:unny* subjects to translate English passives, since SLQZ does not have an English-type passive construction (16):

- (16) R-auhw **bu:unny** buhdy
 hab-eat **person** chicken
 "Chickens are eaten"

Sentences with preverbal *bu:unny* (such as (14)) can only be interpreted as generic statements about people. This is consistent with Kuroda's characterization of generic statements as categorical judgments, and further strengthens the correlation between fronted arguments and subjects of categorical judgments.

In the preceding sections, then, I have shown the following: first (following Kuroda), I assume that categorical judgments emphasize a subject and its relation to its predicate; second, I have shown that focus of arguments in SLQZ is realized by raising of the argument to preverbal position; and finally, focused arguments (syntactic subjects or objects) represent subjects of categorical judgments.

3.1. The Definite Aspect and Thetic Readings

In this section, I will show that thetic readings of clauses with Definite verbs are derived from obligatory verbal focus.

In the overview to this paper, it was shown that Definite verbs disallow preverbal subjects. They disallow other focused elements as well, as seen in (18):

- (17) Laa:a' izhih i-to'oh Gyeihlly ca'rr? (18) *Laa:a' izhih s-to'oh Gyeihlly ca'rr?
 quest tomorrow irr-sell Mike car quest tomorrow def-sell Mike car
 "Will Mike sell the car TOMORROW?" "Will Mike sell the car TOMORROW?"

While temporal adverbials such as "tomorrow" may be focused (and preverbal) in sentences with Irrealis aspect (17), they may not appear preverbally in sentences with Definite aspect (18).

This suggests that Definite verbs themselves necessarily raise to higher positions than do verbs with other aspect markers.

I will thus claim the following:

- Verbs with Definite aspect raise to Focus. This would account for both their incompatibility with other focused elements and their "emphatic future" readings.

Evidence for the presence of verbal focus in Definite sentences comes from the behavior of contrastively focused verbs with other aspects. Like Definite verbs, contrastively focused verbs (19) also disallow preverbal subjects (20):

- (19) B-i:lly Gyeihlly ciity n-gyi'a:a-dya' Gyeihlly
 perf-sing Mike neg subj-dance-neg Mike
 "Mike SANG, not danced"
- (20) *Gyeihlly b-i:lly ciity n-gyi'a:a-dya' Gyeihlly
 Mike perf-sing neg subj-dance-neg Mike
 "Mike SANG, not danced"

This confirms the hypothesis that verbal focus, like argument focus, involves movement to a higher (preverbal) projection. This also reinforces the correlation between verbal focus and the ungrammaticality of preverbal subjects.

This raises the question of how the emphatic future readings of Definite verbs are derived, and why these readings differ from those of other focused verbs, such as those in (19). That is, why does an emphatic reading result when Definite verbs are focused, but only a contrastive reading result when verbs with other aspects are focused?

First, I'll suggest an account for the emphatic readings of Definite verbs. Following Rooth 1992, I assume focus serves to contrast the focused element (in this case, the inflected verb) from a set of other candidates. Thus, the element checked in the focus projection makes a proposition true, while other elements from the set of possible candidates do not.

When a Definite verb (and its lexical and inflectional features) raises to the focus projection, ALL of its features are contrasted against those of other candidates. Thus, the proposition expressed by the sentence could only be true if all of the verb's features are true. For instance, the proposition expressed in (2) (repeated below) is only true if there is a selling event involving a car and Mike, and this event takes place in the future:

- (2) S-to'oh Gyeihlly ca'rr
 def-sell Mike car
 "Mike will (definitely) sell the car"

Thus, of all the possible events that could happen in the world perceived by the speaker, the only one that MUST happen is the event of Mike selling the car sometime in the future.

This is directly analogous to sentences in English in which auxiliaries (that is, the tense and agreement features of the predicate) are focused: these structures also result in emphatic readings (that is, readings in which an event is presupposed to take place):

- (21) Mike DID sell the car
 Mike WILL sell the car
 Mike IS selling the car

Thus, in English focused auxiliary sentences, as in SLQZ Definite sentences, it is not the participants in the event depicted by the verb that are perceived as particularly salient, but

the event as a whole. Thus, sentences with Definite verbs expressthetic judgments. The focus on the tense and agreement features of the predicate draws attention to the predicate—and thus, the event it expresses—rather than to the subject or object of the sentence.

This leaves the question of the derivation of contrastive focus readings on verbs, such as those in (19). In these cases, only the verb itself, not the entire situation in which it appears, is contrasted against alternative candidates. This is directly analogous to the following examples in English:

- (22) Mike will SELL the car (not buy it)
We are SELLING the car (not buying it)

In contrast to the sentences in (21), here only the verb, and not its tense or agreement features, is stressed. While the different focus readings represented by (21) and (22) can be phonologically distinguished in English (since English allows auxiliaries to carry tense and agreement features separately from the main verb) this is not possible in SLQZ; SLQZ verbs obligatorily carry aspect markers (which also encode tense information).

A detailed analysis of the verbal syntax of English versus SLQZ is beyond the scope of this paper. A possibility I will tentatively assume is this: focus features are generated on heads (or XPs) and checked by movement to FocusP. In the case of contrastive focus (as in (19)) the focus features are generated on V, thus causing the verb itself to be contrasted. In the case of Definite verbs, which obligatorily raise to Focus, the focus feature is generated within the Tense/Aspect projections. This causes the Definite verb, which incorporates into Tense and Aspect, to raise into the Focus projection.

This section has shown the following: first, there is empirical evidence that Definite verbs necessarily raise higher than verbs with other aspects, and that this movement involves the preverbal Focus projection. Also, since focus on the verb in Definite sentences reflects speaker attention drawn to an eventuality, rather than to the participants in an event, the obligatory focusing of Definite verbs results in obligatorythetic readings.

4. More Evidence

Further evidence for the obligatorythetic readings of Definite verbs comes from their interaction with certain indefinite expressions. Sentences with Definite aspect cannot, for instance, support specific indefinite readings of "someone/something" or *bu:unny*, "person". This is consistent with Ladusaw's (1994) assertion that weak (non-presupposed) readings of indefinites result from "existential closure due to thethetic mode of judgment":

- (23) Ira'ta' bu:unny ri: jweer nah pehr baall gyaab nnihsyiyah ngaasy
all people go out now but if falls rain later
"Everyone is outside now but if it rains later"

neut-exist/(***def-exist**) bu:unny la:a'ny yu'uh
n-u'uh /***z-u'uh** people in house
there will be people in the house"

In (23), the people that will be in the house are necessarily coreferential with the people outside, and the sentence is ungrammatical with Definite aspect on the verb *ru'uh*, "exist".

In contrast, the Definite form of exist, *zu'uh*, is allowable in contexts in which the indefinite does not refer to some presupposed entity:

- (24) Naahsy ciity tu n-u'uh la:a'ny yu' uh ngaisy **z-u'uh** bu:unny la:a'ny yu'uh
 now neg who neut-exist in house later **def-exist** people in house
 "There's nobody in the house now, but later there will be people in the house"

Similar effects occur with *bu:unny* subjects. *Bu:unny* can occur as the subject of future clauses with either Irrealis or Definite aspect in neutral contexts—that is, those in which no set of people is assumed to exist or not exist. For instance, a sentence such as "People will visit San Lucas next summer" may be uttered in isolation with either an Irrealis or a Definite verb, as seen in (25) and (26):

- (25) **Ch-igueiny** bu:unny bisitaar Sann Luu'c loh beraann
irr-go.to.do person visit San Lucas to summer
 "People will visit San Lucas next summer"
- (26) **Z-igueiny** bu:unny bisitaar Sann Luu'c loh beraann
def-go.to.do person visit San Lucas to summer
 "People will visit San Lucas next summer"

In contexts in which *bu:unny* can only be interpreted as an indefinite, non-specific group of people, only Definite aspect may be used. In (27), for instance, *bu:unny* cannot refer to anyone previously mentioned in the context of the sentence, and the Definite aspect may be used:

- (27) Teebag tu ny-a-dya' wduhbiihahz pehr loh beraann re:e' z-igueiny bu:unny
 neg who subj-come-neg last.year but to summer this **def-go.to.do** person
 "Nobody came last year, but this summer, people will
- bisitaar Sann Luu'c
 visit San Lucas
 visit San Lucas"

In this context, however, the Irrealis form of the verb may not be used:

- (28) *Teebag tu ny-a-dya' uduhbi:ahz pehr loh beraann re:e' **ch-igueiny** bu:unny
 neg who subj-come-neg last.year but to summer this **irr-go.to.do** person
 "Nobody came last year, but this summer, people will
- bisitaar Sann Luu'c
 visit San Lucas
 visit San Lucas"

Conversely, coreferenced *bu:unny* subjects may not appear with Definite verbs. (This is consistent with Kennelly's (1996) proposal that nonspecificity in Turkish is a reflection ofthetic judgments.) For example, a discourse introduced by the sentence in (29), "There are people in Tlacolula now," may be felicitously continued by the sentence in (30), in which the *bu:unny* subject appears as the subject of an Irrealis verb, and is interpreted as "the people who are in Tlacolula now":

- (29) N-u'uh ra bu:unny Ba'c
neut-exist plural person Tlacolula
"There are people in Tlacolula now"
- (30) Loh beraann re:e' ch-iguiny ra bu:unny bisitaar Sann Luu'c
to summer this irr-go.to.do plural person visit San Lucas
"This summer, the people will visit San Lucas"

Sentence (29), however, may not be followed by (31), in which the *bu:unny* subject appears as the subject of a Definite verb:

- (31) # Loh beraann re:e' z-iguiny ra bu:unny bisitaar Sann Luu'c
to summer this def-do plural person visit San Lucas
"This summer, people will visit San Lucas"

In this case, the people who will visit San Lucas cannot be construed as the same people who are now in Tlacolula.

Assuming Ladusaw's and Kennelly's analysis ofthetic judgments as essentially existential, and thus incompatible with specific readings of indefinites, the incompatibility of specific readings of *bu:unny* subjects with Definite verbs can be accounted for: Definite verbs raise to focus, thus drawing attention to events, rather than to their participants. Thus, clauses with Definite verbs representthetic judgments. Asthetic judgments, they are thus incompatible with specific indefinites.

5. Summary

In this paper, I have accounted for the syntactic and semantic constraints on Definite verbs in SLQZ as follows:

- Preverbal subjects in SLQZ are focused, and represent subjects of categorical judgments
- Verbs with Definite aspect raise to focus, which accounts for their emphatic readings and incompatibility with preverbal subjects
- Verbal focus provides a mechanism for derivingthetic readings: focus on the predicate draws attention to the predicate (and the event it describes) rather than to the subject.

References

- Black, Cheryl (1994) *Quiégolani Zapotec Syntax*. UC-Santa Cruz dissertation
- Brennan, Virginia (1995) "Questions in Discourse" *Proceedings of West Coast Conference in Formal Linguistics XIV*, CSLI, Stanford, CA
- Chomsky, Noam (1992) "A Minimalist Program for Syntactic Theory" *MIT Occasional Papers in Linguistics*, 1
- Groenendijk, J., and Stokhof, M. (1984) *The Partitioning of Logical Space*. Doctoral diss.
- Kayne, Richard (1994) *The Antisymmetry of Syntax*. Linguistic Inquiry Monograph 25, MIT Press, Cambridge, MA

- Kennelly, Brenda. (1996) "Nonspecificity = Theticity: An Instantiation of Presentational Focus" To appear in *Proceedings of Student Conference in Linguistics 8*, MIT Working Papers in Linguistics, Cambridge, MA
- Kuroda, S.Y. (1972) "The Categorical and Thetic Judgment: Evidence from Japanese Syntax" *Foundations of Language*
- Ladusaw, William. (1994) "Thetic and Categorical, Stage and Individual, Weak and Strong" In Mandy Harvey and Lynn Santelmann, eds., *Proceedings from Semantics and Linguistic Theory 4*, Ithaca NY: Cornell University, Dept. of Modern Languages and Literatures.
- Lee, Felicia (1994) "Negative Polarity Licensing in Wh-Questions: The Case for Two Licensers" Paper presented the 69th Annual Meeting of the LSA
- Munro, Pamela (1994) "Coreference Phenomena in San Lucas Quiavini Zapotec". Ms., UCLA
- Munro, Pamela, and Felipe Lopez (with Rodrigo Garcia and Olivia Méndez) (1996) *Dicsyonaary X:te:e'n Di:i'zh Sah Sann Luu'c. San Lucas Quiavini Zapotec Dictionary. Diccionario Zapoteco de San Lucas Quiavini*. Ms., UCLA
- Rizzi, Luigi (1995) "The Fine Structure of the Left Periphery." Ms., University of Geneva
- Rooth, Mats (1992) "A Theory of Focus Interpretation " *Natural Language Semantics*, Vol. 1, no.1

Adverbial Quantifiers, Maximal Situations, and “Weak” E-type Pronouns*

Toshiyuki Ogihara
University of Washington

Abstract

This paper argues with von Stechow (1994) and others that adverbs of quantification such as *always* and *usually* are quantifiers over situations, not unselective quantifiers. However, our proposal differs from previous proposals in that it embraces the following ideas: (i) A sentence of the form δ if/when α , β (where δ is a QAdverb) means that δ -many of the maximal situations in which α obtains and throughout which β could conceivably obtain are also β -situations. The domain of quantification for an adverbial quantifier cannot be characterized in terms of minimal situations, however the term *minimality* is defined. Moreover, each situation that serves as a counting unit may not be “extended” into a matrix clause situation. (ii) So-called E-type pronouns always receive a “weak” reading (= Indefinite Lazy Reading for Schubert and Pelletier (1989)) equivalent to an indefinite description, not the standard E-type reading. The proposal defended here is couched in Kratzer’s (1989) situation-theoretic framework, where situations are parts of worlds. We superimpose temporal and spatial ingredients into her system. A sentence of the form if/when p , *always* q is true iff $\{s_1 \mid p$ is true in s_1 and s_1 is a maximal situation such that at any part of s_1 , it is conceivable that p and q is true $\} \subseteq \{s_2 \mid p$ and q is true in $s_2\}$.

1. Introduction

This paper investigates how to determine the domain of quantification for adverbial quantifiers such as *always* and *usually* and argues against the idea that it is determined in terms of minimal situations in which the restrictive clause is true. We propose, instead, that it is determined in terms of *maximal situations* in which the antecedent is true and throughout which it is conceivable that the consequent clause is true. As is well known, since the beginning of the 1980s, the semantics of adverbs of quantification has been a focus of attention among formal semanticists. Based upon Lewis’s (1975) idea that such adverbs can bind multiple variables and hence are “unselective quantifiers,” Kamp (1981) and Heim (1982) independently developed a theory referred to as Discourse Representation Theory (henceforth DRT for short). DRT was used to account for donkey sentences, among other things. Although the DRT approach to natural language semantics has produced many interesting research results, it has many empirical problems, which cast doubt upon the validity of the basic idea that DRT embraces: adverbial quantifiers are unselective quantifiers in that they bind unlimited number of free variables that occur within the restrictive clause.

One major problem with DRT is the so-called proportion problem. It is illustrated by example (1) (Kadmon 1987):

- (1) Most women who own a cat are happy.

The classical DRT analysis predicts that (1) is true in the following scenario: among the ten cat-owning women, one owns 100 cats and is happy, whereas the other women own one cat each and are unhappy. This is because 100 woman-cat pairs verify the condition whereas only nine woman-cat pairs fail to do so. However, (1) is intuitively false in the circumstance just described. In terms of DRT, this means that we must modify the theory in such a way that *most* behaves like a selective quantifier that only binds the variable associated with *women* in (1). The desired interpretation is obtained by the traditional generalized quantifier approach to the semantics of NPs coupled with the existential quantifier analysis of indefinite NPs (e.g., Montague 1973). The same problem arises with the conditional variant of donkey sentences. Consider example (2).

- (2) If a woman owns a cat, she is usually happy.

In the original DRT analysis, (2) is understood to have the same truth conditions as (1). This is intuitively incorrect. Bäuerle and Egli (1985) suggest that we can account for examples like (1) and (2) on the basis of the following generalizations: (i) when an indefinite NP in the restrictive clause of a quantifier is not anaphorically linked to a pronoun in the nuclear scope of the quantifier, the indefinite NP is interpreted as existentially quantifying; (ii) when an indefinite NP in the restrictive clause is anaphorically linked to a pronoun in the nuclear scope, they are understood as occurrences of the same variable and are caught by the adverbial quantifier.

However, this generalization fails when we look at examples like (3a–b):

- (3) a. Drummers mostly live in crowded dormitories. But if a drummer lives in an APARTMENT COMPLEX, it is usually half empty.
b. If a man has a quarter in his pocket, he usually puts it in the parking meter.

(3a) is due to Heim (1990), and (3b) is discussed by Schubert and Pelletier (1989). (3a) shows that despite the fact that *an apartment complex* in the *if*-clause is anaphorically linked to the pronoun *it* in the matrix clause, the sentence can be interpreted in such a way that *usually* quantifies over the set of drummers. That is, (3a) can receive an interpretation symbolized in (4a). In this case, the pronoun *it* receives an E-type interpretation because *it* is paraphrased as “the apartment that *x* lives in.” (3b) is also problematic. Its natural interpretation only requires that each man who has a quarter put at least one quarter in the parking meter, as indicated in (4b).

- (4) a. usually_x ∃y[drummer(*x*), apartment complex(*y*), *x* lives in *y*][the unique apartment complex *z* in which *x* lives is half empty] (an E-type reading of *it*)
b. usually_x ∃y[man(*x*), quarter(*y*), *x* has *y* in *x*'s pocket]∃z[quarter(*z*), *x* has *z* in *x*'s pocket, *x* puts *z* in the parking meter] (an indefinite lazy reading of *it*, which does not require the presence of a unique quarter for each man.)

Schubert and Pelletier (1989) refer to this reading as an indefinite lazy reading. This poses a problem for any variant of the E-type analysis as long as it seeks to

preserve the uniqueness presupposition associated with (so-called) E-type pronouns. We will take up this matter in more detail in the next section.

In order to solve the empirical problems associated with the classical DRT analysis, some researchers (e.g., Berman 1987, Heim 1990, von Stechow 1994) have proposed situation-based analyses of donkey sentences and some related phenomena. The formal theory of situation assumed in these proposals is that of Kratzer (1989). The ontology of Kratzer's theory is given in (5).

- (5) Ontology of Kratzer's Situation Theory
- S a set, the set of possible situations
 - A a subset of S , the set of possible individuals
 - \leq a partial ordering on S , with at least one additional condition: for all $s \in S$ there is a unique $s' \in S$ such that $s \leq s'$ and for all $s'' \in S$, if $s' \leq s''$, then $s'' = s'$.
 - $\wp(S)$ the power set of S , the set of propositions
 - W a subset of S , the set of maximal elements with respect to \leq . W is the set of possible worlds. For all $s \in S$, let w_s be the maximal element s is related to by \leq .

The idea underlying the situation-based proposals such as Berman (1987), Heim (1990) and von Stechow (1994) is that quantificational adverbs quantify over one type of object only, i.e., situations. To correctly restrict the domain of quantification for adverbial quantifiers, these proposals hold that they quantify over minimal situations of the relevant sort. When the sentence in question has an overt restrictive clause, for example an *if*-clause, the domain is claimed to be the following set: $\{s \mid s \text{ is a minimal situation such that } \textit{if}\text{-clause is true in } s \text{ and } s \in C\}$, where C is the set of situations provided by the previous context. For example, in (2) *usually* quantifies over $\{s \mid s \text{ is a minimal situation such that there is a cat-owning woman } x \text{ in } s \text{ and } s \in C\}$. This theory assumes that we can somehow pragmatically select minimal situations that contain a cat-owning women as "counting units" for *usually*.

(6a) contains a new indefinite NP in the nuclear scope and is analyzed as in (6b).

- (6) a. If a farmer owns a donkey, he usually sells it to a merchant.
 b. usually, [s is a minimal situation in which a farmer owns a donkey]
 $\exists s' [s \text{ is part of } s', \text{ the unique farmer } x \text{ who owns a donkey } y \text{ in } s \text{ sells the donkey } x \text{ owns in } s \text{ to a merchant in } s']$

As indicated in (6b), one is allowed to extend each "minimal situation" to find a situation in which the consequent is true. However, the proposal cannot account for examples like (7) which involve the problem of indistinguishable participants (Heim 1990, von Stechow 1994).

- (7) If a man has the same name as another man, he usually avoids addressing him by name.

(7) is problematic under any situation-based proposal that adopts an E-type analysis of pronouns. Note that any (minimal) situation in which the antecedent is true must contain two men. Therefore there is no unique man in such a situation. The theory predicts that the pronouns *he* and *him* have no denotation and therefore (7) is

uninterpretable. However, (7) in fact receives a perfectly coherent interpretation. This is a serious problem for proposals that are based upon minimal situations.

Next, note that most sentences that involve adverbial quantifiers are purely or partially time-sensitive. Nevertheless, previous situation-based proposals have largely ignored temporal matters. Kratzer (1989) abstracts away from temporal issues and pretends that each situation is atemporal for the sake of simplicity. This simplified situation theory does not have enough machinery to fully account for the semantics of adverbial quantifiers. We tentatively propose the following extension of Kratzer's situation theory that incorporates times: Let I be the set of intervals defined in the usual way. Let τ be a function (called the "temporal trace function") from situations to intervals. We also posit partial orders $<_t$ (strict temporal precedence) and \sqsubseteq_t (subinterval relation) on the set of intervals. For example, $\tau(s_1) <_t \tau(s_2)$ says ' s_1 temporally precedes s_2 ', and $\tau(s_1) \sqsubseteq_t \tau(s_2)$ means ' $\tau(s_1)$ is a subinterval of $\tau(s_2)$ '.

- (8) a. If a woman buys a sage plant here, she usually buys eight others along with it.
 b. Before John visits Mary, he always calls her.
 c. If/When a farmer owns a donkey, he usually beats it.

The selection of each relevant situation for (8a) clearly involves temporal considerations. To determine the statistical tendency of how people make purchases of sage plants at a particular store, it is necessary to conduct a survey over a period of time. Thus, it is not possible to avoid the question of time in order to obtain the right interpretation of (8a). Intuitively, multiple sage-plant-buying events can constitute one buying situation only if they are temporally close to one another. For example, if one person buys sage plants on nine different days, these nine buying events normally cannot be grouped together as one buying situation. Obviously, one cannot predetermine how close the events have to be to qualify as one counting unit (situation, for our purposes) for *usually*. We must decide on a case-by-case basis, taking into consideration such factors as people's intentions. Intuitively, (8a) is true iff in most maximal situations s such that s serves as a "sage-plant-buying-situation" and a woman buys at least one sage plant in s , the woman buys nine of them in s . However, this is not what the standard situation-based analysis predicts because (8a) involves the proportion problem. Let us present one concrete case. Assume that there are eleven women who buy some sage plants on various occasions. A woman buys twenty sage plants on one occasion, and the other ten women buy a sage plant each. Let us assume that they are the only sage-plant-buying events that are relevant to the evaluation of (8a). Intuitively, *usually* quantifies over the eleven occasions or "situations." Since only one woman buys at least nine sage plants, the sentence is intuitively false. However, there are thirty minimal situations in which a woman buys a sage plant, and twenty of them can be extended to a larger situation in which the same woman buys eight other sage plants.¹ The standard situation-based account therefore fails to do justice to our intuitions associated with example (8a). We shall discuss a solution to this problem in the next section.

It is not obvious how to extend a minimality-based proposal to account for examples involving *before*-clauses or *after*-clauses, but it seems reasonable to propose for (8b) the truth condition described in (9).

- (9) $\{s \mid s \text{ is a minimal situation in which John visits Mary}\} \subseteq \{s_1 \mid \exists s_2 \exists s_3 [s_1 \supseteq s_2 \ \& \ \tau(s_3) \subseteq \tau(s_2) \ \& \ \text{John calls Mary in } s_3 \ \& \ \tau(s_3) <_t \tau(s_1)]\}$

(9) says that every minimal situation in which John visits Mary can be extended to a larger situation in which John calls Mary before he visits her. Unfortunately, this proposal faces a problem Partee (1984) discusses: it predicts that a single event of John's calling Mary that precedes all events of John's visiting Mary is enough to make (8b) true because each minimal situation associated with the restrictive clause may be extended indefinitely until it incorporates the single event of John's phone call to Mary. This is clearly an incorrect prediction.

It turns out that problems associated with time are more pervasive than they appear at first. No proposal based upon minimal situations can handle the classical donkey sentence (8c), at least not straightforwardly. According to the standard situation-theoretic account, the domain of quantification for *usually* is the following set: $\{s \mid s \text{ is a minimal situation such that a farmer owns a donkey in } s \text{ and } s \in C\}$. If we disregard C , this results in the wrong prediction about the truth condition for (8c). Note that the restrictive clause in (8c) is a stative sentence. It is usually assumed (e.g., Bennett and Partee 1972) that stative sentences can be defined in terms of the subinterval property as shown in (10a–b).

- (10) a. ϕ is said to have the **subinterval property** iff for any interval t if ϕ is true at t , then ϕ is true at all the subintervals of t .
 b. ϕ is a stative sentence iff it has the subinterval property.

Given this assumption, we are obliged to conclude that if there is an interval t at which a stative sentence ϕ is true then there are infinitely many sub-intervals of t at which ϕ is true. If we assume that time is dense, there is no minimal interval at which ϕ is true.² Given the mapping relations between intervals and situations posited above, we are obliged to conclude that there is no minimal situation in which ϕ is true. This is a problem for a theory based on minimal situations because it predicts that there is no minimal situation in which the restrictive clause is true. One obvious way out is to rely on the contextually salient situations indicated by C . That is, we can say in principle that the set of contextually salient situations C filters out situations that are too small and selects the right ones that correspond to maximal stretches of a man's having a donkey. However, (8c) is interpretable even when it is uttered out of context. When the context does not restrict the domain of quantification in any way, each minimal situation is presumably determined by the meaning of the restrictive clause alone. Therefore, the fact that a minimal situation in which the restrictive clause is true is non-existent in cases like (8c) significantly weakens the main claim made by the proposals based upon minimal situations.

2. A Proposal Based upon Maximal Situations

Previous situation-based proposals have thus failed to make empirically accurate predictions. However, the simplicity of their approach is very appealing. In many cases, quantificational adverbs simply quantify over times, and it would be nice if we could extend this basic function of these expressions to cover a wider range of cases. The proposal I will advance incorporates an important idea adopted in situation-theory-based proposals, namely that adverbial quantifiers quantify over situations. However, I make the following claims, which are in disagreement with

the previous proposals made within a situation-based theory: (i) the domain of quantification for an adverbial quantifier cannot be determined in terms of minimal situations in which the *if/when* clause is true; (ii) the correct truth condition cannot be determined by allowing the original situation to be extended into a nuclear scope situation. We must set up the system in such a way that the domain of quantification consists of maximal situations of some sort so that actual or potential main-clause events can occur within them.

As a first step toward an improved proposal, let us clarify the relations between situations and spatio-temporal regions. Kratzer's (1989) situation theory is designed to account for what she refers to as the "lumping relation" between propositions. Assuming that Paula painted apples and bananas yesterday evening, Kratzer observes that the fact Paula painted a still life somehow "includes" the fact that she painted apples in the actual world. Put differently, whatever makes (11a) true in the actual world also makes (11b) true.

- (11) a. Paula painted a still life.
b. Paula painted apples.

In Kratzer's terms, (11a) lumps (11b) in the actual world. Given this intuition about the lumping relation between (11a) and (11b), Kratzer's situation theory characterizes it as follows: the minimal situation in which (11a) is true includes the minimal situation in which (11b) is true.

In presenting her situation theory, Kratzer (1989) carefully notes that situations cannot be identified with spatio-temporal regions. She points out that if she is hungry and tired at the same time, the minimal space-time chunk in which she is hungry would also be the minimal space-time chunk in which she is tired. Therefore, if situations were just spatio-temporal regions, these two propositions would be expected to lump each other. However, this goes against our intuition. Therefore, Kratzer posits situations as primitive entities. However, the basic intuition about the lumping relation clearly comes from the temporal or spatial inclusion relation between two eventualities. For example, the intuition about lumping Kratzer discusses regarding (11a–b) concerns the temporal (and perhaps spatial) inclusion relation between the two "events" in question. Therefore, although two distinct situations may share the same spatio-temporal region, if two situations are ordered via the "part of" relation \leq , we can assume that this is replicated in the temporal domain or in the spatial domain. Given these assumptions, I propose the following. For the purpose of this paper, let us assume that the model contains a set of spaces, each element of which is a set of spatial points that are "connected." On this assumption, the intuitive notion of "spatial sub-part of" can be encoded in terms of the subset relation between two spaces. We posit the function π from situations to spaces, which intuitively indicates the spatial trace of a situation. Then we posit the following mapping relation between situations and spatio-temporal regions that they occupy.

- (12) Functions from situations to times and spaces:
For any situations s and s' , if $s < s'$ then (i) $\tau(s) \subset \tau(s')$ and $\pi(s) \subset \pi(s')$, or
(ii) $\tau(s) \subset \tau(s')$ and $\pi(s) = \pi(s')$, or (iii) $\tau(s) = \tau(s')$ and $\pi(s) \subset \pi(s')$.
Note: " τ " is a function from situation to intervals; π is a function from situations to spaces. " \subset " indicates proper subset.

(12) encodes the aforementioned idea, namely that if a situation s is included in a situation s' , then this is replicated in the temporal domain or in the spatial domain (or both). This means that the following possibilities are disallowed: (i) $s < s'$ and $\tau(s') \subset \tau(s)$; (ii) $s < s'$ and $\pi(s') \subset \pi(s)$. I think these assumptions are intuitive and reasonable. In the tradition of temporal semantics, we say that the proposition p is true at an interval t when the time slice t is just enough to support the truth of p . The idea is different from the minimal interval at which p is true. For example, if John is in his room from 10 to 11, then (13) is assumed to be true at every subinterval of $\{t \mid 10 \leq t \leq 11\}$.

(13) John is in his room.

Although a similar notion in the spatial domain is not commonly discussed in the literature, we can assume that the same technical notion also applies to the case of space. That is, I assume that a proposition is true “at” a space and can also be true “in” a space.

With this preliminary discussion in mind, we can now characterize the concept of “truth at a situation” as in (14).

(14) “Truth at a time and at a space” is a primitive notion related to “truth in a situation” in the following way: For any proposition p and for any situation s , if p is true at $\tau(s)$ and at $\pi(s)$, we say p is true at s and p is true in all situations s' such that $s \leq s'$.

The concept of “truth of some proposition p at a situation” is not the same as the concept of “minimal situation in which p is true.” For example, if John stays in his room from 10 to 11, then (13) is true at a situation s such that $\tau(s)$ equals this one hour interval. However, this is not a minimal interval at which (13) is true because it is a stative sentence and has the subinterval property. Put informally, the main ideas contained in our proposal can be stated as in (15).

(15) Our proposal: (i) On the assumption that an adverb of quantification is a quantifier over situations, the situations with respect to which the restrictive clause is evaluated must be the same as those with respect to which the nuclear scope is evaluated. (ii) The domain of quantification for the adverbial quantifier *always* in a sentence of the form *if p, then always q* is the set of maximal situations in which p obtains and throughout which q 's being true is conceivable. (iii) Unbound pronouns are always interpreted as if they are indefinite descriptions.

It is arguable that the right notion is neither “minimal situation” nor “maximal situation.” That is, an adverbial quantifier quantifies over a set of situations such that the size of each such situation totally depends upon the context of use. However, as mentioned earlier some sentences that involve an adverb of quantification are uttered out of context and yet interpretable. This means that we somehow determine the “counting units” correctly from the content of the sentence alone. I argue that the correct counting units are characterized as maximal situations in which the adverbial clause is true and throughout which the truth of the matrix clause is conceivable.

In order to interpret unbound variables as disguised indefinite descriptions, I posit the following rules.

- (16) a. Assign a numerical index to each NP.
 b. Adjoin each non-pronominal NP to the minimal S that contains it.
 c. Copy the restrictive clause to the nuclear scope. Schematically, any sentence of the form δ , *if/when* α , *then* β (where δ is an adverb of quantification) converts into δ , *if* α , *then* [α and β].
 d. Existentially close both the restrictive clause and the nuclear scope.

This proposal is based upon a preliminary analysis presented by Chierchia (1992), who refers it to Heim (personal communication).^{3, 4} We shall see that the rules (16a–d) make the right predictions with regard to the examples we considered so far. Our implementation is different from Chierchia's in that it is situation-based and all indefinite NPs are existentially quantifying. In Chierchia's formulation, some indefinite NPs are singled out by the rule of topic selection and get bound by an adverbial quantifier. It is not clear how time is dealt with in Chierchia's (1992) proposal. Since an adverbial quantifier is a selective quantifier for Chierchia, it may or may not bind a time variable that occurs in the restrictive clause. Either way, we will encounter a problem. If a time variable is caught by the quantifier, we run into the problem pointed out above. That is, as soon as we find one instantiation of the time variable that makes the restrictive clause true, there are an infinite number of them. On the other hand, if a time variable is not caught by the quantifier, then it is caught by the existential quantifier. This also gives us the wrong result. For example, if one and the same woman got pregnant twice, these two pregnancies would not count as two counting units with regard to (17).

- (17) If a woman gets pregnant, she usually sees a doctor immediately.

Thus, we must deal with the problem associated with time anyway, and I believe that our proposal is a step in the right direction.

Thus, we adopt (16) to account for the semantics of donkey pronouns. On the basis of (16), any sentence of the form (18a) is transformed into a structure given in (18b).⁵

- (18) a. If p , then always q .
 b. always, if $\exists[p]$, then $\exists[p$ and $q]$

(18b) is not enough to predict the right truth conditions for (18a). To obtain the right domain of quantification for *always* (and other adverbial quantifiers), (18b) is further modified as in (19a), which yields the truth conditions described informally in (19b).

- (19) a. always_s , if s is a maximal situation such that $\exists[p]$ in s & $\forall s_1[s_1 \leq s \rightarrow [\Diamond [\exists[p$ and $q]$ at $s_1]]]$ & $s \in C$, then $\exists[p$ and $q]$ in s
 Note: " $\Diamond p$ " reads 'it is conceivable that p .'
 b. $\{s \mid \exists[p]$ is true in s and s is a maximal situation such that at any sub-situation of s it is conceivable that $\exists[p$ and $q]$ is true and $s \in C\}$
 $\subseteq \{s \mid \exists[p$ and $q]$ is true in $s\}$

The semantics of \Diamond is not explicitly provided here, and we will discuss it below.

Armed with the proposal just presented, we shall re-examine (2), repeated here as (20).

(20) If a woman owns a cat, she is usually happy.

(20) is transformed as in (21).

- (21) a. If a woman₁ owns a cat₂, she₁ is usually happy.
 b. usually, if $[s_{\text{NP a woman}}]_1[s_{\text{NP a cat}}]_2[s e_1 \text{ owns } e_2]$, $[s_{\text{NP a woman}}]_1[s_{\text{NP a cat}}]_2[s e_1 \text{ owns } e_2]$ and she₁ is happy
 c. usually, if $\exists_{1,2}[[s_{\text{NP a woman}}]_1[s_{\text{NP a cat}}]_2[s e_1 \text{ owns } e_2]]$, $\exists_{1,2}[[s_{\text{NP a woman}}]_1[s_{\text{NP a cat}}]_2[s e_1 \text{ owns } e_2]]$ and she₁ is happy
 d. usually_s, if $\exists_{1,2}[[s_{\text{NP a woman}}]_1[s_{\text{NP a cat}}]_2[s e_1 \text{ owns } e_2]]$ in s & $\forall s_1[s_1 \leq s \rightarrow \{\Diamond[\exists_{1,2}[[s_{\text{NP a woman}}]_1[s_{\text{NP a cat}}]_2[s e_1 \text{ owns } e_2]]]$ and she₁ is happy] at s_1] & $s \in C$, $\exists_{1,2}[[s_{\text{NP a woman}}]_1[s_{\text{NP a cat}}]_2[s e_1 \text{ owns } e_2]]$ and she₁ is happy] in s

On the basis of (21d), the domain of quantification for *usually* is obtained as follows: $\{s \mid \text{a woman owns a cat in } s \text{ and } s \text{ is a maximal situation such that at any sub-situation of } s \text{ it is conceivable that a woman who owns a cat is happy and } s \in C\}$. In this case, it seems reasonable to use the set of maximal situations at which a woman owns a cat as the domain of quantification. If the same woman owns a cat at two discontinuous situations (i.e., at different intervals), we must evaluate these situations separately. This possibility is usually not considered in conjunction with this example, but this is in fact the right empirical generalization.

Let us see how (16a–c) apply to example (3b), repeated here as (22).

- (22) If a man has a quarter in his pocket, he usually puts it in the parking meter.
 (23) a. if $[a \text{ man}]_1$ has $[a \text{ quarter}]_2$ in his₁ pocket, he₁ usually puts it₂ in the parking meter.
 b. usually, if $[s[a \text{ man}]]_1[s[a \text{ quarter}]]_2[s e_1 \text{ has } e_2 \text{ in his}_1 \text{ pocket}]$, $[s[a \text{ man}]]_1[s[a \text{ quarter}]]_2[s e_1 \text{ has } e_2 \text{ in his}_1 \text{ pocket}]$ and he₁ puts it₂ in the parking meter
 c. usually, if $\exists_{1,2}[[s[a \text{ man}]]_1[s[a \text{ quarter}]]_2[s e_1 \text{ has } e_2 \text{ in his}_1 \text{ pocket}]]$, $\exists_{1,2}[[s[a \text{ man}]]_1[s[a \text{ quarter}]]_2[s e_1 \text{ has } e_2 \text{ in his}_1 \text{ pocket}]]$ and he₁ puts it₂ in the parking meter
 d. usually_s, if s is a maximal situation such that $\exists_{1,2}[[s[a \text{ man}]]_1[s[a \text{ quarter}]]_2[s e_1 \text{ has } e_2 \text{ in his}_1 \text{ pocket}]]$ in s & $\forall s_1[s_1 \leq s \rightarrow \Diamond[[\exists_{1,2}[[s[a \text{ man}]]_1[s[a \text{ quarter}]]_2[s e_1 \text{ has } e_2 \text{ in his}_1 \text{ pocket}]]]$ and he₁ puts it₂ in the parking meter] at s_1] & $s \in C$, $\exists_{1,2}[[s[a \text{ man}]]_1[s[a \text{ quarter}]]_2[s e_1 \text{ has } e_2 \text{ in his}_1 \text{ pocket}]]$ and he₁ puts it₂ in the parking meter in s

According to (23d), (22) receives the following interpretation: most maximal situations in which a man has a quarter and throughout which it is conceivable that a man who has a quarter puts a quarter he has in the parking meter are situations in which a man who has a quarter puts a quarter he has in the parking meter. On this proposal, we can represent our intuition about how to find the domain of quantification in terms of maximal situations. For example, with regard to times, we can easily identify maximal intervals in terms of whether some person is at the parking meter continuously. The idea is that if a situation does not contain a person who has a quarter standing near the parking meter, we assume that it is not conceivable that someone puts a quarter in the parking meter in this situation. This

enables us to do justice to our intuition that in (3b) *usually* quantifies over maximal situations s throughout which a man who has a quarter in his pocket stands near the parking meter.

Consider now (7), repeated here as (24).

- (24) If a man has the same name as another man, he usually avoids addressing him by name.

It is syntactically analyzed as in (25).

- (25) a. if [a man]₁ has the same name as [another man]₂, he₁ usually avoids addressing him₂ by name
 b. usually, if [s[a man]₁][s[another man]₂][s e₁ has the same name as e₂]], [s[a man]₁][s[another man]₂][s e₁ has the same name as e₂]] and he₁ avoids addressing him₂ by name
 c. usually, if $\exists_{1,2}[s[a \text{ man}]_1[s[\text{another man}]_2[s e_1 \text{ has the same name as } e_2]]]$, $\exists_{1,2}[[s[a \text{ man}]_1[s[\text{another man}]_2[s e_1 \text{ has the same name as } e_2]]]$ and he₁ avoids addressing him₂ by name]
 d. usually, if s is a maximal situation such that $\exists_{1,2}[s[a \text{ man}]_1[s[\text{another man}]_2[s e_1 \text{ has the same name as } e_2]]]$ in s & $\forall_{s_1}[s_1 \leq s \rightarrow \diamond [[\exists_{1,2}[[s[a \text{ man}]_1[s[\text{another man}]_2[s e_1 \text{ has the same name as } e_2]]]]]$ and he₁ avoids addressing him₂ by name]] in s_1] & $s \in C$, $\exists_{1,2}[[s[a \text{ man}]_1[s[\text{another man}]_2[s e_1 \text{ has the same name as } e_2]]]$ and he₁ avoids addressing him₂ by name] in s

On the basis of (25d), one can arrive at the following truth conditions for the sentence (24): most maximal situations in which a man has the same name as another man and throughout which it is conceivable that a man x who has the same name as another man y avoids addressing y by name are also situations in which a man x who has the same name as another man y avoids addressing y by name. The domain of quantification in this case consists of maximal situations throughout which two men have the same name and they know each other because in such situations it is conceivable that they avoid addressing each other by name.

Let us now consider some complex and crucial examples. (8a–c) are repeated here as (26a–c).

- (26) a. If a woman buys a sage plant here, she usually buys eight others along with it.
 b. Before John visits Mary, he always calls her.
 c. If/When a farmer owns a donkey, he usually beats it.

Our proposal analyzes (26a) as in (27).

- (27) a. usually, if $\exists_{1,2}[s[a \text{ woman}]_1[s[a \text{ sage plant}]_2[s \text{ e}_1 \text{ buys e}_2 \text{ here}]]]$, $\exists_{1,2}[[s[a \text{ woman}]_1[s[a \text{ sage plant}]_2[s \text{ e}_1 \text{ buys e}_2 \text{ here}]]]$ and she₁ buys eight others along with it₂
 b. usually, if s is a maximal situation such that $\exists_{1,2}[s[a \text{ woman}]_1[s[a \text{ sage plant}]_2[s \text{ e}_1 \text{ buys e}_2 \text{ here}]]]$ in s & $\forall s_1[s_1 \leq s \rightarrow \Diamond[[\exists_{1,2}[[s[a \text{ woman}]_1[s[a \text{ sage plant}]_2[s \text{ e}_1 \text{ buys e}_2 \text{ here}]]]]]$ and she₁ buys eight others along with it₂] at s_1] & $s \in C$, then $\exists_{1,2}[[s[a \text{ woman}]_1[s[a \text{ sage plant}]_2[s \text{ e}_1 \text{ buys e}_2 \text{ here}]]]$ and she₁ buys eight others along with it₂] in s

(27b) receives the following interpretation: most maximal situations in which a woman buys a sage plant here and throughout which it is conceivable that a woman buys some sage plant x here and buys eight others along with x are situations in which a woman buys some sage plant z here and also buys eight others along with z . This provides the right truth condition for (26a).

(26b) is an example that makes a different point. As pointed out earlier, previous situation-based proposals allow for the possibility that each restrictive clause situation is expanded into a consequent situation, and this yields the wrong result. In our proposal, each counting unit is determined in part by the main clause, and it is used to evaluate the nuclear scope of an adverbial quantifier. This makes the right predictions. I assume here that a *before*-clause is used to characterize situations that are located immediately before a situation at which the *before*-clause is true. The domain of quantification for *always* is the following set of situations: $\{s \mid \text{John visits Mary immediately before } s \text{ and } s \text{ is a maximal situation such that at any sub-situation of } s \text{ it is conceivable that John calls Mary before he visits her}\}$. Thus, (26b) is true iff $\{s \mid \text{John visits Mary immediately before } s \text{ and } s \text{ is a maximal situation such that at any sub-situation of } s \text{ it is conceivable that John calls Mary before he visits her and } s \in C\}$ is a subset of $\{s \mid (\text{John visits Mary immediately before } s \text{ and) John calls Mary in } s\}$.

Lastly, let us discuss how to characterize the concept of “conceivably true,” which is admittedly a fuzzy concept. We are talking about some type of possibility here, and I offer the characterization in (28) as a first approximation.

- (28) For any situation s , $\Diamond[\phi \text{ at } s]$ is true iff there is a proposition ψ such that the truth of ψ significantly increases the chances of ϕ 's being true and ψ is true at s .

The idea is that any situation s that could be used as a “counting unit” is a maximal s such that some relevant proposition is true throughout s . For example, with regard to (22) we can use a maximal situation throughout which (29) is true.

- (29) A man is standing in front of the parking meter with a quarter in his pocket

It seems natural to assume that the truth of this proposition significantly increases the chances of satisfying the condition given in the nuclear scope. Although this characterization is still very informal and rough, I hope it helps to make my idea clear to the reader.

3. Remaining Problems and Issues

It has been pointed out in the literature that examples like (30) (Heim 1990) receive “unselective binding” readings. In fact, it is difficult to interpret (30) in any other way.

- (30) Most people who owned a slave owned his children and grandchildren too.

I believe that this type of reading is not an independent interpretation assigned to the sentence by the semantic component but is forced upon us by some pragmatic factors. Our proposal only assigns a weaker interpretation to (30): Most people who own a slave owned the children and grandchildren of a slave they owned. I contend that this is in fact the only interpretation that (30) receives. Since it is very odd to assume that one owns several slaves but does not own the children and grandchildren of all of them, one tends to “assign” an unselective reading to (30). To see that this view is plausible, consider (31), which has the same structure as (30) in the relevant respects.

- (31) Most people who use a credit card for purchases use it for cash advances too.

(31) does not have an “unselective” interpretation, at least not obligatorily. The most natural interpretation of (31) is the reading predicted by our proposal. That is, it is enough for someone to use a credit card for purchases and to use a different credit card for cash advances to satisfy the condition described by the sentence. Thus, I think we can conclude that the weaker interpretation is in fact the only interpretation available to unbound pronouns linked to indefinite NPs.

I believe that the main idea incorporated in our proposal can be recast in terms of a different framework. For example, it should be possible to reinterpret situations in terms of tuples that involve (at least) times, spaces, and objects. However, the central claims made in this paper are valid all the same because any framework must deal with time. That is, the claims made in this paper regarding situations translates into the following claims about times: (i) the domain of quantification for an adverbial quantifier should be determined in terms of maximal intervals, rather than minimal intervals; (ii) each interval that serves as a “counting unit” for an adverbial quantifier may not be extended when the nuclear scope is evaluated.

Endnotes

* I would like to thank Bill Ladusaw, Virginia Brennan, Kai von Stechow, Yuki Matsuda, and the audiences at Kyushu University and the University of California at Santa Cruz for comments and suggestions. All errors are my own.

¹ Obviously, we can let the context filter out those situations that are too small to serve as counting units for *usually*, but there is no principled way of predicting the right interpretation for any given case.

² Time is dense iff for any two distinct times t_1 and t_2 such that $t_1 <_t t_2$, there is a time t_3 such that $t_1 <_t t_3 <_t t_2$, where $<_t$ is used to indicate strict temporal precedence.

³ Chierchia's (1992) official proposal is couched in a dynamic semantic system and differs from the preliminary proposal considered here.

⁴ Kratzer (1988) shows that this analysis can be seen as the E-type analysis of pronouns coupled with the Heimian proposal about definite descriptions.

⁵ Adverbial quantifiers are proposed to create a tripartite structure.

Bibliography

- Bäuerle, Rainer and Urs Egli (1985) "Anapher, Nominalphrase und Eselssätze," Papier 105 des Sonderforschungsbereichs 99, Universität Konstanz.
- Berman, Steve (1987) "Situation-based Semantics for Adverbs of Quantification," in J. Blevins and A. Vainikka (eds.), *University of Massachusetts Occasional Papers* 12, University of Massachusetts, Amherst.
- Bennett, Michael and Barbara Partee. (1972) 'Toward the Logic of Tense and Aspect in English'. Distributed by Indiana University Linguistics Club, Bloomington.
- Chierchia, Gennaro (1992) "Anaphora and Dynamic Binding," *Linguistics and Philosophy* 15, 111–183.
- Heim, Irene (1982) *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. dissertation, University of Massachusetts, Amherst.
- Heim, Irene (1990) "E-type Pronouns and Donkey Anaphora," *Linguistics and Philosophy* 13, 137–177.
- Kamp, Hans (1981) "A Theory of Truth and Semantic Representation," in J. Groenendijk et al. (eds.), *Truth, Interpretation and Information*, Foris, Dordrecht.
- Kadmon, Nirit (1987) On Unique and Non-unique Reference and Asymmetric Quantification, Ph.D. dissertation, University of Massachusetts, Amherst.
- Kratzer, Angelika (1988) "Stage-Level and Individual-Level Predicates," in M. Krifka (ed.), *Genericity in Natural Language*, 247–284. SNS-Bericht 88-42, University of Tübingen.
- Kratzer, Angelika (1989) "An Investigation of the Lumps of Thought," *Linguistics and Philosophy* 12, 607–653.
- Lewis, David (1975) "Adverbs of quantification," in Ed. Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge University Press, Cambridge.
- Montague, Richard. (1973) 'The Proper Treatment of Quantification in Ordinary English', in Jaakko Hintikka, J. M. E. Moravcsik and Patrick Suppes (eds.), *Approaches to Natural Language, Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, D. Reidel, Dordrecht.
- Partee, Barbara (1984) "Nominal and Temporal Anaphora," *Linguistics and Philosophy* 7, 243–286.
- Schubert, Lenhart K. and Francis J. Pelletier (1989) "Generically Speaking, or, Using Discourse Representation Theory to Interpret Generics," in Gennaro Chierchia, et al. (eds.), *Properties, Types and Meaning, Volume II: Semantic Issues*, Kluwer, Dordrecht.

von Stechow, Kai (1994) *Restrictions on Quantifier Domains*, Ph.D. dissertation,
University of Massachusetts, Amherst.

Department of Linguistics
University of Washington
Box 354340
Seattle, WA 98195-4340

ogihara@u.washington.edu
<http://weber.u.washington.edu/~ogihara>

When the best isn't good enough

Phonologically-conditioned ungrammaticality in Optimality Theory

Cemil Orhan Orgun
UC Davis

Ronald Sprouse
UC Berkeley

1. Introduction*

One of the major insights of Optimality Theory (Prince and Smolensky 1993) is that grammatical constraints are ranked and violable. An infinite set of candidate forms is evaluated by these ranked constraints, and the winning candidate in effect is a compromise between the potentially conflicting demands imposed by these constraints. A question that has been only rarely addressed in the OT literature is how UNGRAMMATICALITY arises if all constraints are violable in principle and constraint violation does not entail ungrammaticality. In this paper, we point to some shortcomings of the only existing proposal to deal with ungrammaticality in OT, the special constraint MPARSE (Prince and Smolensky 1993), and propose a restructuring of EVAL.

We propose the addition of another constraint component called CONTROL,[†] which contains only those inviolable constraints that may cause ungrammaticality (rather than repair). The winning candidate from EVAL, the usual ranked and violable constraint component, is submitted to CONTROL. If this candidate satisfies all the constraints in CONTROL, it is a grammatical output. If it violates a constraint in CONTROL, no grammatical output is possible. Our approach is not only empirically superior to MPARSE, but it also makes a clear distinction between two kinds of inviolable constraints. Inviolable constraints in EVAL (those that outrank all potentially conflicting constraints) cause repairs or block otherwise general alternations; inviolable constraints in CONTROL cause ungrammaticality, never repair.

2. MPARSE (Prince and Smolensky 1993, Raffelsiefen 1996)

Prince and Smolensky (1993) propose that the output of GEN always includes a special candidate called the Null Parse, which has no phonetic realization and is

* Larry Hyman and Sharon Inkelas have provided much valuable input to this paper. Earlier versions of this paper were presented at the TREND conference at Stanford, and at WECOL. We thank both audiences, especially Junko Itô, Armin Mester and Jaye Padgett. We are also grateful to Sharon Rose for detailed comments on an earlier version of this paper.

[†] This paper was written during Sunday homebrewing sessions over the course of several months. Originally, we entertained the brilliant idea of expressing some deep connection between homebrewing and linguistic research by using the term SPARGE for our new constraint component. This was a beautiful idea, we thought, as out of the sparge comes a wort! However, like many good ideas, this one too has in the end lost to commonsense, which dictated that we use a term like CONTROL, which, while certainly much more mundane, is likely, we felt, to be more accessible to a wider audience.

stipulated to satisfy all well-formedness and faithfulness constraints.¹ By definition, the Null Parse violates only the special constraint MPARSE, which no other candidate violates.² Ranking a phonological constraint C above MPARSE is equivalent to declaring it inviolable; any candidate that violates C is worse than the Null Parse. In (1), the Null Parse emerges as the winning candidate because all other candidates violate constraint C, which outranks MPARSE.

(1) C » MPARSE; ungrammaticality preferred over violation of C

	Constraint C	MPARSE
Candidate A	*!	
☞ Null Parse		*

In the next section, we illustrate the use of MPARSE to deal with ungrammaticality by summarizing Raffelsiefen's (1996) analysis of morphological gaps in English *-ize* suffixation.

3. English *-ize* formation (Raffelsiefen 1996)

Raffelsiefen (1996) claims that the English verbalizing suffix *-ize* can be attached productively to adjectives with non-final stress, but not to adjectives with final stress.

(2)	<u>Non-final stress</u>		<u>Final stress</u>	
	rándom	rándomize	corrúpt	*corruptize
	vápor	váporize	obscéne	*obscene-ize
	átom	átomize	secúre	*secure-ize

Raffelsiefen argues that ungrammaticality in the case of final-stressed adjectives results from an irresolvable conflict between two constraints, *CLASH and IDENT.

- (3) *CLASH: Two adjacent stressed syllables are prohibited.
 (4) IDENT: The stem of the derived word must be identical to the base.
 (i.e., no stress shift)

Both *CLASH and IDENT are ranked above MPARSE. Any candidate that violates either *CLASH or IDENT is worse than the Null Parse. Since it is impossible to simultaneously satisfy both *CLASH and IDENT for an input form with final stress, the Null Parse emerges as the winning candidate.

¹ P&S assume that the Null Parse is identical to the input but fails to be inserted into the morphological structure. In this paper, we abstract away from any specific approach to morphology and concentrate on the empirical question of how to identify ungrammatical outputs.

² For Raffelsiefen the candidate that violates MPARSE is identical to the input and may violate other constraints.

(5) Stem with final stress \Rightarrow no grammatical output³

	kərɒpt-áyʒ	IDENT	*CLASH	MPARSE
	kórəptàʒ	*!		
	kərɒptàʒ		*!	
☞	Null Parse			*

When the input stem has non-final stress, however, it is possible to satisfy both *CLASH and IDENT. Therefore, a grammatical output is possible for adjectives with non-final stress.

(6) Stem with non-final stress \Rightarrow grammatical output

	rændəm-ayz	IDENT	*CLASH	MPARSE
☞	rændəmàʒ			
	Null Parse			*!

We have seen that Prince and Smolensky's MPARSE approach handles ungrammaticality by ranking their special constraint MPARSE below other grammatical constraints. Whenever violation of one of those higher-ranked constraints is unavoidable, the MPARSE-violating Null Parse emerges as the winning candidate. No grammatical output is possible in such cases.

While MPARSE works for many cases of ungrammaticality, we will see that it is unable to handle all cases. In the remaining sections, we will present data from Turkish, Tagalog, and Tiene where an MPARSE account would predict a grammatical output where there is in fact no grammatical output.

4. Challenges to MPARSE

In this section, we show that there are cases of ungrammaticality in which the ungrammatical candidate could be repaired by violating a constraint independently known to be violable in other (grammatical) output forms in the language. If constraint C is violable, it has to be ranked below MPARSE. In that case, MPARSE cannot force ungrammaticality since violation of MPARSE (by the Null Parse) is

³ Raffelsiefen actually gives a phonological representation to the Null Parse, which is identical to the phonological input form, with the affixes unattached to their stems. This amounts to saying that the affix is unable to attach to the stem. In this approach to MPARSE, the Null Parse does violate other constraints. This assumption makes it harder (in comparison with P&S's approach) for MPARSE to be the constraint causing ungrammaticality in other cases, as the Null Parse may be ruled out by higher-ranked constraints. As we will see, the problem with MPARSE is that it fails to rule out some ungrammatical candidates. Since Raffelsiefen's use of MPARSE is weaker than P&S's it can only do worse with regard to these problems, not better. Accordingly, we assume P&S's approach in the rest of this paper.

more serious than violation of C (by another candidate). The Null Parse therefore cannot be the winning candidate.

4.1 Turkish

Our first challenge to MPARSE comes from Turkish, where subminimal forms are not repaired by epenthesis, even though epenthesis is found elsewhere in the language.

4.1.1 Minimal size condition

As Itô and Hankamer (1989) and Inkelas and Orgun (1995) observe, some speakers of Turkish impose a disyllabic minimal size constraint on suffixed forms.

(7)	Root	Suffixed form ($\sigma\sigma$ min)		
	sol ^y	'note G'	sol ^y -üm	'my G'
	do:	'note C'	*do:-m	'my C'

Ungrammatical monosyllabic forms are not augmented by epenthesis, as shown in (8), where epenthetic segments are enclosed in boxes:

- (8) Repair by epenthesis is not possible
- a) *do: yu m
 - b) *do: u m
 - c) *i do: m
 - d) *do: m u

The failure of ungrammatical forms to be augmented by epenthesis suggests that constraints barring epenthesis must outrank MPARSE, allowing the Null Parse to win. The relevant grammatical constraints are DEPRT, which disallows epenthesis of segments, and LEX \approx PR and FTBIN, which together require each form to contain a disyllabic foot.

- (9) DEPRT Do not insert segments (McCarthy and Prince 1995)
- (10) LEX \approx PR, FTBIN Every word must contain a disyllabic foot (Prince and Smolensky 1993)

Since violating these constraints in order to create a grammatical output form is not possible, they must all outrank MPARSE (11):

- (11) DEPRT, LEX \approx PR, FTBIN » MPARSE

The tableau in (12) shows how this ranking accounts for the failure of subminimal forms to be repaired by epenthesis—all epenthetic candidates violate a constraint ranked above MPARSE, as does the subminimal candidate. The Null Parse therefore emerges as the winning candidate:

(12) Tableau for input /do: - m/

/do:-m/	DEPRT	LEX≈PR, FtBIN	MPARSE
do:m		*!	
do:yum			
do:um			
ido:m	*!*		
do:mu	*!		
☞ Null Parse			

While this analysis accounts for the current set of data, it suffers from a crucial flaw: ranking DEPRT above MPARSE implies that epenthesis is never possible. However, epenthesis is in fact allowed to avoid vowel hiatus or illicit coda clusters in suffixed forms:

- (13) /araba + a/ → araba^ya 'car-dative'
 /it + m/ → itⁱm 'my dog'

If MPARSE were to allow epenthesis, it must outrank DEPRT. This ranking is shown in (14), where the candidate with epenthesis is the winner.

(14) Tableau for input /it - m/ ⇒ different ranking required

/it-m/	MPARSE	DEPRT	LEX≈PR, FtBIN
☞ itim			
Null Parse	*!		

This results in a ranking paradox. The analysis in (14) requires that MPARSE outrank DEPRT, whereas (12) requires that DEPRT outrank MPARSE. An MPARSE account must either overgenerate, allowing subminimal forms to be augmented by epenthesis, or undergenerate, disallowing epenthesis into clusters.

In the following section, we propose a solution to this dilemma.

4.1.2 Solution: CONTROL

Even when there is no grammatical output, speakers often have judgments about what the output would have been if a grammatical output were possible. Our proposal takes these intuitions as a starting point in developing an empirically superior alternative to MPARSE. Specifically, we propose that the ranked

constraint component *EVAL* always produces a winning candidate, an optimal form with respect to the given constraint ranking. In order to deal with ungrammaticality, we introduce a new evaluation component, *CONTROL*, which is a set of inviolable constraints that winning candidates from *EVAL* must satisfy in order to be accepted as grammatical output forms.⁴ This proposal is based on the important but not immediately obvious observation that natural language grammars contain two different types of inviolable constraints. The first type is commonly discussed in OT—constraints that force violation of lower-ranked constraints but are never violated themselves.⁵ The second kind, which has not received as much attention in the literature, causes ungrammaticality but never repair. Placing both types of constraints in *EVAL* leads to ranking paradoxes, as in Turkish. Placing the second (ungrammaticality-causing) type of inviolable constraint into the new component *CONTROL* avoids these paradoxes. In order to be grammatical, an output must satisfy two conditions: (i) it must be the optimal candidate chosen by *EVAL*; and (ii) it must satisfy all constraints in *CONTROL*.

- (15) Conditions required for grammatical output. The output must
- i) be the optimal candidate chosen by *EVAL*;
 - ii) satisfy all constraints in *CONTROL*.

The resolution to the Turkish problem is in (16). Since the minimality conditions on derived surface forms in Turkish never force augmentation of a subminimal form, or any other kind of repair, minimality constraints belong in *CONTROL*, not *EVAL*. In (16) the winning candidate of *EVAL* is *do:m*, which violates neither of the constraints in *EVAL*. However, when this winning candidate is submitted to evaluation in *CONTROL*, it fails to satisfy the minimality conditions and is therefore ungrammatical, as indicated by the \bowtie symbol.

(16) Input /do:-m/

<i>EVAL</i>		
/do:-m/	*VV	DEPRT
\bowtie do:m		
do:um	*!	
do:yum		*!*

⁴ *CONTROL* resembles Halle's (1973) notion of the Filter in that a portion of the grammar is allowed to overgenerate and ungrammatical forms are filtered out, but the formal mechanism is quite different. *CONTROL* is more restricted in its use, and, unlike Halle's Filter, cannot alter the phonological, morphosyntactic or semantic properties of its inputs.

⁵ The Turkish constraint causing epenthesis in (13) is an example of an inviolable constraint that always triggers repair.

CONTROL	
	LEX \approx PR, FtBIN
✗	*!

The winning candidate from EVAL is submitted to CONTROL strictly for grammaticality judgments. Unlike EVAL, CONTROL evaluates a single form, and therefore does not choose between candidates; it can only declare the single winning candidate from EVAL grammatical or ungrammatical. Therefore, no repair is possible to satisfy constraints in CONTROL. If the winning candidate from EVAL violates a constraint in CONTROL, ungrammaticality results.

The MPARSE account was confounded by the fact that epenthesis is possible elsewhere in Turkish even though it is not used to augment subminimal forms. Our solution is not subject to this difficulty. By ranking CODACOND above DEPRT in EVAL, for example, we can account for the fact that epenthesis is used in order to avoid illicit coda clusters (17):

(17) Input /it + m/

EVAL		
/it + m/	CODACOND	DEPRT
✗ itim		*
itim		*!*

CONTROL	LEX \approx PR, FtBIN
✓ itim	*!

This example illustrates the important difference between two types of inviolable constraints that our proposal seeks to capture. CODACOND is never violated in Turkish. It is always obeyed, even at the expense of violating lower-ranking faithfulness constraints in EVAL. The prosodic minimality condition is also never violated, but forms that violate it are judged ungrammatical and never repaired. We account for this by placing the minimal size constraints in CONTROL rather than EVAL. CODACOND, on the other hand, must be in EVAL, since it demonstrably interacts with other constraints in EVAL.

4.1.3 Constraining GEN: a failed attempt to save MPARSE

Inviolable constraints have sometimes been suggested to be a part of GEN (e.g., McCarthy and Prince 1995). Usually these proposals have been intended to capture cross-linguistic universals, but it is also worth considering whether appropriately constraining GEN might resolve the Turkish ranking paradox without requiring the use of SPARGE. Suppose GEN were prohibited from creating derived output candidates of less than two syllables. While this move successfully removes the subminimal candidates, it incorrectly predicts that the epenthetic

candidate in (20) will win over the Null Parse. MPARSE must still outrank DEPRT since epenthesis is allowed in order to prevent CODACOND violations. Consequently, the candidate that violates only DEPRT is preferred over the Null Parse.⁶

(18) Tableau for input /do: - m/

/do:-m/	MPARSE	DEPRT
do:um		
Null Parse	*!	

Constraining GEN, therefore, cannot be the right approach.

4.2 Tagalog *-um-* infixation

Infixation of the verbal marker *-um-* poses a similar challenge to an MPARSE analysis. Following M&P's (1993) analysis of Tagalog infixation, we assume that *-um-* aligns to the left edge of the word, but it infixes into consonant-initial roots in order to yield superior syllable structure (i.e., to avoid codas). In (19) *-um-* prefixes to a vowel-initial stem and infixes when the stem begins with a consonant or consonant cluster.

- (19) abot um-abot 'to reach for'
 sulat s-um-ulat 'to write'
 gradwet gr-um-adwet 'to graduate' (French 1988)

(20) ALIGN(L, *um*) The morpheme *-um-* is located at the left edge; is a prefix.
 (This is Edgemost in M&P and P&S)

(21) NOCODA Syllables are open.

Although our analysis of *-um-* infixation is in the same spirit as M&P's, we will amend it slightly to deal with additional facts that we have elicited from our Tagalog consultants. For some speakers, NOCODA cannot be the constraint that drives infixation. For these speakers, infixation of *-um-* into CC-initial stems does not necessarily yield the smallest possible number of NOCODA violations. Instead, *-um-* infixation improves syllable structure by avoiding onsetless syllables whenever possible. Thus, *-um-* may be infixated after the first consonant in a CC-initial stem (22) ([Avery, 1995 #1]):

⁶ We cannot constrain GEN to prevent it from performing epenthesis, which, as we have already seen, is possible in Turkish.

- (22) gradwet grumadwet ~ gumradwet 'to graduate'
 plantsa plumantsa ~ pumlantsa 'to iron'
 preno pumreno ~ prumeno 'to brake'

Following Anttila 1995, we assume that variation results from crucial non-ranking of constraints with respect to one another. In Tagalog, we claim that NOCODA and ALIGN are not crucially ranked with respect to each other. In (23), *-um-* is added to the CC-initial stem *plantsa*. If NOCODA is favored over ALIGN, *plumantsa* emerges as the winner. If ALIGN is favored over NOCODA, however, ONSET forces infixation by only one segment, and *pumlantsa* is the winner.

- (23) ONSET » NOCODA, ALIGN ⇒ variable infixation

/um + plantsa/	ONSET	NOCODA	ALIGN
☞ <u>p</u> umlantsa		**	
☞ plumantsa		*	
umplantsa	*!	**	

In any analysis using MPARSE, MPARSE must outrank ONSET, ALIGN, and NOCODA, as affixation of *-um-* results in grammatical forms violating all three of these constraints.

- (24) MPARSE » ONSET, ALIGN, NOCODA ⇒ prefixation

um + abot	MPARSE	ONSET	NOCODA	ALIGN
☞ <u>u</u> mabot		*	*	
Null Parse	*!			
abumot		*	*	u!

Having established that MPARSE outranks ALIGN, we will now present cases of ungrammaticality that could have been prevented by additional alignment violations. These cases show that MPARSE is not a viable option for dealing with ungrammaticality in Tagalog infixation: ALIGN is known to be violable (even multiply violable, as in *gr-um-adwet*) in Tagalog. Given that, it must be ranked below MPARSE. As such, ungrammaticality cannot be caused by ALIGN. The relevant restriction is that *-um-* cannot attach to *w-* or *m-*initial stems, a fact that we attribute to an OCP-related constraint. OCP violations are not avoided by violating ALIGN as would be expected in an MPARSE analysis since MPARSE must outrank ALIGN; instead, they result in ungrammaticality. The fact that ALIGN is not violated to avoid ungrammaticality implies that it must outrank MPARSE. We therefore encounter a ranking paradox.

Three pieces of evidence illustrate the ungrammaticality of *-um-* with *m-* and *w-*initial stems:

1. Distribution:

-um- never occurs with native words beginning with /m/ or /w/: "It may be noted that *-um-* does not occur with bases beginning with /m/ or /w/." (Schachter and Otnes 1972; 292)

2. Variable infixation of *-um-* is not possible with /Cw/-initial native stems:

gwapo	gumwapo ~ *gwumapo	'become handsome'
sweti	sumweti ~ *swumeti	'become sweaty'

3. *m*-initial loans cannot take *-um-*:⁷

foggy →	fumafagi na	'it's foggy now'
cloudy →	kumaklawdi na	'it's cloudy now'
misty →	*mumimisti na	'it's misty now'

In the next section, we show that MPARSE cannot deal with these data.

4.2.1 Failure of MPARSE

The OCP violations in Tagalog could in principle be avoided by hyper-infixation of *-um-* into the stem, but this never occurs, a fact that could not be captured with an MPARSE analysis.

(25) OCP-*um* **m-um*, **w-um*

This OCP constraint must outrank MPARSE in order to cause ungrammaticality. This ranking alone is not sufficient to derive ungrammaticality, however. In order for MPARSE to cause ungrammaticality, it must be outranked by at least two constraints that potentially conflict with each other. When satisfying one of those constraints that outrank MPARSE entails violating another, the Null Parse will emerge as the winner. In Tagalog, it is obvious that the relevant constraint that conflicts with OCP is ALIGN, for violating ALIGN would have been a way to satisfy OCP. Accounting for ungrammaticality therefore requires ranking both OCP and ALIGN above MPARSE. Yet, MPARSE must outrank ALIGN since alignment violations *are* tolerated in *-um-* infixation. This ranking incorrectly predicts that further alignment violations should be allowed in order to prevent OCP violations (26):

⁷ These examples involve reduplication in addition to infixation. *tawag* 'call' → *tumatawag* is a grammatical output of this particular morphological construction.

(26) MPARSE cannot rule out ungrammatical candidate with hyper-infixation

um + RED + misti	OCP	MPARSE	ONSET	NoCODA	ALIGN
mumimisti	*!				
ummimisti			*!		
mimistumi					
Null Parse		*!			

We have seen that an MPARSE analysis of Tagalog *-um*-infixation encounters a fatal ranking paradox. In the following section, we illustrate that a principled solution using CONTROL is readily available.

4.2.2 Solution using CONTROL

Since OCP-*um* causes ungrammaticality rather than repair, it must be in CONTROL, not EVAL. The correct result is then obtained as shown in (27):

(27) The winning candidate from EVAL is ruled out by OCP-*um* in CONTROL

um + RED + misti	ONSET	NoCODA	ALIGN
mumimisti		*	m
ummimisti	*	*	
mimistumi		*	mimist

CONTROL	OCP- <i>um</i>
mumimisti	*!

Our CONTROL account succeeds where MPARSE failed, thanks to its proper separation of inviolable constraints that interact with the rest of the system to cause repair from those that do not interact with the grammar.

4.3 Tiene

Our next example comes from Tiene. In Tiene, deletion is required in order to avoid violating a constraint on stem shape that prohibits CVCVC stems with coronal consonants in the onsets of both the final and penultimate syllables, which we refer to as STEM SHAPE. However, a constraint on stem *size* that rules out stems containing more than three syllable leads to ungrammaticality rather than repair by deletion.

Deletion triggered by STEM SHAPE is illustrated in (28). The data come from Ellington 1977. Our analysis closely follows that of Hyman (1996).

(28)	Deletion:				
	mata	'go away'	maasa	'cause to go away'	
	bóta	'give birth'	bóose	'deliver (child)'	
	kɔɔ	'become tired'	kɔɔsɔ	'tire' (tr.)	
	pala	'arrive'	paasa	'cause to arrive'	
	píina	'be black'	píise	'blacken'	
	banya	'be judged'	baasa	'cause to be judged'	

In these forms, the causative suffix (Proto Bantu *-is) is added to stems whose last consonant is a coronal. STEM_{SHAPE} violation is avoided by deleting the coronal stem consonant (see Hyman 1996 and Hyman and Inkelas 1997 for details).

Having established that deletion is allowed by the grammar of Tiene, we show that a constraint violation that could have been avoided by deleting a consonant instead leads to ungrammaticality. The construction of interest is definitive aspect formation.

For a disyllabic stem, reduplicating the last stem syllable (29) forms the definitive aspect:⁸

(29)	yɔbɔ	'bathe'	yɔbɔbɔ	'bathe thoroughly'
	mata	'go away'	matata	'go away once and for all'
	yaka	'believe'	yakaka	'believe once and for all'
	lɔŋɔ	'load'	lɔŋɔŋɔ	'load once and for all'

For stems containing more than two syllables, no morphologically expressed definitive form is possible. A periphrastic form must be used instead.⁹

(30)	kótoba	'chase'	*kótobaba	*kóobaba
	vúteke	'come back'	*vútekeke	*véekeke
	binema	'sleep'	*binemama	*beemama
	panama	'frighten'	*panamama	*paaamama

The constraint responsible for ungrammaticality is STEM_{SIZE}, which restricts stems to a maximum size of three syllables.¹⁰ The fact that deletion cannot salvage

⁸ The fact that STEM_{SHAPE} is violated in these forms shows that it must be outranked by constraints on base-reduplicant identity. Since STEM_{SHAPE} in turn outranks constraints against deletion, the situation is hopeless for M_{PARSE}, which would need constraints barring deletion to be inviolable in order to cause ungrammaticality.

⁹ The periphrastic definitive aspect is formed by adding *nkó msto* to the conjugated verb in the neutral aspect (Ellington 1977:93)

STEMSIZE violations implies that, in an MPARSE account, DEPRT outranks MPARSE. However, this incorrectly rules out deletion in the causative forms in (28). This is an irresolvable ranking paradox.

CONTROL circumvents this problem entirely: STEM SIZE is in CONTROL, as are all constraints that cause ungrammaticality rather than repair. Therefore, nothing motivates deletion in the definitive aspect forms in (30). The winning candidate from EVAL contains four syllables, and the STEM SIZE constraint in CONTROL rules it out.

(31)	/mata + RED/	MAX	RED= σ
	☞ matata		
	maata	*!	

	/panama + RED/	MAX	RED= σ
	☞ panamama		
	paamama	*!	

	CONTROL	STEM SIZE
	✓ matata	
	✗ panamama	*!

Our Tiene analysis yields a significant insight into the failure of MPARSE. In a rule-based account, environments, targets, and repairs are bundled into a single package. Thus, one does not expect a rule to apply outside its intended environment. As pointed out by Prince and Smolensky and McCarthy and Prince, this packaging prevents rule-based accounts from capturing interesting generalizations, as there are cases within and across languages where a single target may be reached by various paths depending on the input form. Optimality Theory provides a more satisfactory approach by decoupling wellformedness targets from the operations that allow a language to reach them. The actual way in which the wellformedness targets are reached (or fail to be reached) emerges from the interaction of grammatical wellformedness and faithfulness constraints.

While this architecture gives rise to aesthetically pleasing accounts, it also gives rise to an interesting potential problem: a given repair is sometimes available in a particular environment, but not in others. For example, in Tiene, a consonant may be deleted under pressure from the OCP, but not under pressure from the maximal size condition. In a rule-based account, this situation could be handled by building the OCP into the deletion rule's environment. In Optimality

¹⁰ The three syllable maximum could be stated as Ft + σ to avoid counting. Alternatively, as Inkelas and Hyman suggest, a maximal size constraint of two syllables (one foot) could be imposed on the *core stem*, that is, the verb root plus all tense/aspect suffixes, excluding the final vowel morpheme.

Theory, the process must be decoupled from the target. This makes it possible for the process to apply in unexpected environments.

Our approach using CONTROL allows us to once again decouple certain constraints from possible repair procedures. This is empirically required, since there are cases where repairs used to avoid violating some constraints are not resorted to when other constraints are violated, giving rise to ungrammaticality instead of repair. However, our approach is still more restricted than a rule-based one, since we predict that an inviolable constraint that fails to interact with one grammatical constraint must in fact fail to interact with *any* grammatical constraint.

5. Conclusion

In the three languages we have discussed, MPARSE fails to rule out some ungrammatical forms. In particular, it fails when the following conditions obtain:

- a constraint C_1 is independently known to be violable in some grammatical output forms;
- a different constraint C_2 is never violated;
- violation of C_2 can be avoided by violating C_1 .

When these conditions hold, MPARSE predicts that there will be a grammatical output that violates C_1 and satisfies C_2 . However, we have shown that there are cases in which precisely these conditions obtain, yet there is no grammatical output.

Turkish does not augment subminimal forms, even though its grammar allows epenthesis to satisfy syllable structure constraints.

In Tagalog hyper-infixation is not used in order to prevent unwanted *m-um-* and *w-um-* sequences even though multiple alignment violations are allowed in regular infixation.

In Tiene, deletion is not possible to avoid violations of STEM SIZE, even though deletion *is* available to avoid violations of STEM SHAPE.

We have proposed adding a second constraint component, CONTROL, that consists only of inviolable constraints that cause ungrammaticality rather than repair. Any grammatical form has to be the winner of EVAL and satisfy all the constraints in CONTROL.

We maintain that our approach to ungrammaticality is superior to MPARSE on both empirical and theoretical grounds:

CONTROL is able to account for ungrammaticality of forms that could apparently be salvaged by (additional) violation of a low-ranked constraint.

MPARSE is vague in that it stipulates that a Null Parse of any one morpheme makes the whole string grammatically uninterpretable. Our approach makes explicit the reasons for ruling out entire strings. Inputs to CONTROL are always whole strings and are evaluated that way.

Finally, CONTROL offers a principled account of the difference between two types of inviolable constraint: those that cause ungrammaticality are in CONTROL. Those that cause repair or block alternations are in EVAL, and outrank all conflicting constraints.

References

- Anttila, Arto. 1995. Deriving variation from grammar: a study of Finnish genitives. Manuscript, Stanford University.
- Avery, Peter and Greg Lamontagne. 1995. Infixation <and metathesis> in Tagalog. Canadian Linguistic Association.
- Ellington, John Ernest. 1977. *Aspects of the Tiene language*. Doctoral dissertation, University of Wisconsin-Madison.
- Hyman, Larry. 1996. Prosodic constraints on morphology: the strange case of Tiene. Linguistics department colloquium, University of California, Berkeley.
- Inkelas, Sharon and Cemil Orhan Orgun. 1995. Level ordering and economy in the lexical phonology of Turkish. *Language* 71(4): 763-793.
- Itô, Junko and Jorge Hankamer. 1989. Notes on monosyllabism in Turkish. In Junko Itô and Jeff Runner (eds.), *Phonology at Santa Cruz*. Santa Cruz: Linguistics research center, University of California. 61-69.
- McCarthy, John and Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. Amherst, MA: Graduate Linguistic Student Association. 249-384.
- Prince, Alan and Paul Smolensky. 1993. Optimality Theory: constraint interaction in generative grammar. Manuscript, Rutgers University and the University of Colorado at Boulder.
- Raffelsiefen, Renate. 1996. Gaps in word formation. In Ursula Kleinhenz (ed.), *Interfaces in phonology*. Berlin: Akademie Verlag. 194-209.
- Schachter, Paul and F. T. Otones. 1972. *Tagalog reference grammar*. Berkeley: University of California Press.

Durational contrasts and the Iambic/Trochaic Law

Anthi Revithiadou and Ruben van de Vijver

Holland Institute of generative Linguistics

Rijksuniversiteit Leiden, Vrije Universiteit Amsterdam, the Netherlands

1. Introduction

Basing himself on psychological experiments on rhythmic grouping, Hayes proposed the Iambic/Trochaic Law.¹ The Law gives a principled formulation of the findings of these experiments and makes a generalization to stress theory possible. The tendency found by psychologists was that people perceive elements contrasting in intensity as being leftheaded, while elements contrasting in duration are perceived as being righthheaded. Hayes assumed that these results reflected not just a preference, but rather they illustrated a wellformedness principle regarding rhythm. The Iambic/Trochaic Law, now, asserts that, in case of an intensity contrast, the preferred grouping has initial prominence, while in case of a durational contrast, the most prominent element is the last one. Consider the Iambic/Trochaic Law (1).

- (1) Iambic/Trochaic Law (Hayes 1995:80).
 a. Elements contrasting in intensity naturally form a trochee.
 b. Elements contrasting in duration naturally form an iamb.

Linguistically, this means that the stressed syllable in iambic feet is longer than the unstressed syllable, whereas in trochaic feet syllables are of equal duration. Hayes mentions that, in iambic feet, stressed syllables must be 1.5 to 2 times as long as unstressed syllables.² Unfortunately, Hayes never refers to this ratio in any analysis of lengthening at all. If the length contrast in iambic feet is enhanced by deletion of the unstressed syllable this ratio is rather difficult to interpret (for example in Eastern Ojibwa). One of the things argued in this paper is that such radical length contrasts can be seen in trochaic languages as well (cf. Proto-Nordic and Mathimathi).

For completeness sake, we would like to mention that (1a) says that elements contrasting in *intensity* form leftheaded groupings (i.e. trochees). Because this point has never been elaborated upon we will refrain from elaborating this point here.

Under a very strict reading this Law says that if there is a length contrast in a language, the language will have iambs, while languages lacking such a contrast are trochaic. This is an extreme and fruitless interpretation of Hayes's Law, since there are many languages that have a weight- or a length contrast which are nevertheless trochaic (Kager 1993). A more promising

¹ We would like to thank Harry van der Hulst, Martin Honcoop, Jan Kooij, Sergio Menuzzi and Micheal Redford for valuable comments. Furthermore, we would like to thank NWO and HIL who made this research possible, financially.

² This raises the obvious question how *1.5 to 2 times as long* should be interpreted phonologically.

approach, one overtly advocated in Hayes (1995), is to say that languages which deploy iambs as their stress feet are susceptible to enhancing durational contrasts in their feet. This enhancement consists of either lengthening of the stressed syllable or reduction of the unstressed syllable in the foot.

As instances of this more tolerant interpretation, Hayes mentions the lengthening of stressed vowels in Hixkaryana (1995:206) and the deletion of unstressed vowels in Eastern Ojibwa (1995:216). These languages have a stress pattern that is analyzed iambically by Hayes. The rightheadedness of these languages implies an enhancement of durational contrasts given the more lighthearted interpretation of the Iambic/Trochaic Law.

Examples of trochaic languages exhibiting the same phenomena—lengthening if stressed, reduction if unstressed—are excluded in the light of the Law. Yet, in the present paper we like to present empirical evidence which instantiates exactly this unexpected situation. More specifically, in some *trochaic* languages unstressed vowels are deleted. Furthermore, there are trochaic languages in which the syllable bearing stress is prolonged. Consequently, stressed vowels become much longer than their unstressed counterparts.

The reduction of syllables in the weak position of the foot is found, among others, in the *trochaic* languages Northern Greek dialects, Tiberian Hebrew, (Byelo)Russian and Pashto. All these languages can be analyzed with trochaic feet, as we will show later in section 2. In these languages weak positions in a foot are reduced or eliminated. The lengthening of stressed syllables in *trochaic* languages will be established with data from Modern Greek, Dutch, German and Swedish. These languages are uncontroversially trochaic (Malikouti-Drachman & Drachman 1988 for Greek; van der Hulst 1984, Kager 1989 for Dutch; Wiese 1996 for German; Riad 1992 for Swedish).

In short, we will argue that the interaction between stress and length should be reconsidered. All stressed syllables are lengthened and all unstressed syllables are shortened. The proposal put forward in this paper is summarized in (2).

(2) *Stress and Length Principle.*

For every σ , L : $\acute{\sigma}$ > σ

Legenda: L = language, $\acute{\sigma}$ = stressed syllable, > = is longer than, σ = unstressed syllable

In section 2, some case studies are presented involving reduction processes in trochaic languages. Section 3 continues with case studies involving lengthening of the stressed syllable in trochaic languages. Section 4 provides an overview of the main ideas explored in this paper.

2 Shortening Modern Greek

In the trochaic language Modern Greek unstressed syllables are reduced, as we will argue in this section. But since the presence of secondary stresses is not an uncontroversial issue, we will first present evidence which support its existence.

Arvaniti (1991) has not found any phonetic evidence for rhythmic stress in Greek. The short duration, low amplitude integral and the low and falling F0 of syllables which are thought to have rhythmic stress makes these syllables very similar to unstressed syllables (Arvaniti 1991:91). She also examined the hypothesis that alternating rhythm may arise due to vowel reductions. In this way, the unreduced vowels become more prominent, and thus rhythmically strong.

Greek has a five vowel system /a, e, i, o, u/ all of which have the same phonological weight. However, only the high vowels /i, u/ can be considerably reduced in unstressed positions. The phenomenon has been studied both phonologically (Theophanopoulou-Kontou 1973) and phonetically (Dauer 1980). Dauer observes that phonetic context and stress pattern play a major part in regulating high vowel reduction in Greek. A high vowel is more susceptible to reduction when it is preceded and/or followed by voiceless consonant, in particular /s/, and *when in a syllable immediately following a stressed one*. In contrast, a vowel surrounded by voiced consonants or immediately preceding a stressed syllable is least likely to be reduced. From an acoustic point of view, Dauer observes that there are various stages of high vowel reduction, ranging from very short high vowels (up to 30 ms) with full formant structure following nasals and laterals, to voiceless high vowels (Dauer 1980:18).

Arvaniti's third experiment (1991:53-74) supports Dauer's findings. In the 24 tokens of the word /akuSTikan/ 'they were heard', there is a total of 8 reductions of /u/ in the environment in which reduction could create an alternating pattern, and a total of only 6 reductions in the word /akustiKA/ 'headphones', in which /u/ can carry rhythmic stress. Dauer's observation that vowels are reduced more often in post-stressed syllables is borne out by /i/ of /aKUstikan/ 'they were heard'³ which is reduced in a total of 15 tokens out of 24. In contrast, in /akustiKA/ 'headphones', where /i/ precedes the stressed syllable, it is reduced in only 7 out of 24 tokens. Arvaniti, based on the results of her experiment, casts doubt on the importance of vowel reduction in creating rhythmic patterns and argues against the inclusion of rhythmic stress in phonological representations. We maintain that such a claim is partly unjustifiable. First of all, the specific phenomenon has never been subject of thorough laboratory investigation. Second, Greek /u/ reduction is sociolinguistically more stigmatized than /i/ reduction and, therefore sophisticated style of Greek bans it as dialectal or indicative of lower social

³ The stress patterns of /aKUstikan/ and /akuSTikan/ 'they were heard' are equally common in Modern Greek.

environment. Finally, both Dauer's and Arvaniti's surveys found that reduction does occur at least within the domain of the (trochaic) foot which bears primary accent. Thus, perhaps high vowel reduction is not very strong argument for claiming that rhythmic stress exists in Greek but has a significant importance for the main proposal of this paper, more specifically, that durational contrast do occur in trochaic languages.

Northern Greek Dialects

The Northern dialects (ND) are spoken in the northern part of mainland and islands in Greece. Under this term are grouped dialects which exhibit the phenomenon of high vowel loss and mid vowel raising in unstressed positions before consonants. A dialect which possesses these phenomena will exploit them to the full. While much fluctuation is unpredictable, certain factors may be singled out as militating against the occurrence of at least loss. Some are essentially phonological while others are metrical in nature. The lack of a systematic study for each one of these dialects restrain us from enumerate here the specific segmental or metrical conditions which impede loss in all of these dialects. Thus, we will concentrate on facts from the dialect of Siatista (SD) which has been meticulously analysed in Margariti-Roga (1985).

Stress in ND, and in SD in particular, follows to a great extent the accentual behavior of Standard Modern Greek. Thus, it is dependent on lexical marks and it is trochaic. As expected, accent shifts interact with both loss and raising. The following examples picture the interaction between stress and reduction. The bracketed forms represent the corresponding Standard Greek words, consonants with an accent are palatalized:

- (3) a. *é-i alternation*
 (pésu) /péso/ (épi)sa /épesa/ 'fall'
 b. *i-0 alternation*
 sku(písu) /skupiso/ (skúpsa) /skúpisa/ 'sweep'
 c. *ó-u alternation*
 pla(kósu) /plakóso/ (pláku)sa /plákosa/ 'crash upon'
 d. *u-0 alternation*
 a(kúsu) /akúso/ (ák)sa /ákusa/ 'hear'

The asymmetry between the set of stressed and unstressed vowels could be expressed in terms of phonological theories such as dependency or government phonology as simplification of complex segments. Such a theory is given here for completeness sake. Van der Hulst and Drescher (1995) employ a theory of segmental representation which ranks the vowels occurring in weak syllabic positions as less complex, or in our terms *reduced*. Under a dependency or government phonology approach vowels are represented as combinations of basic particles or elements. Some representations are set out in (4):

In this case unstressed /e/ sounds in the pronunciation of some people as [i] (de Bray 1980:188), e.g. *teljehráma* [tiljihráma] 'telegram'.

Similarly, in Russian stressed syllables are always very strongly emphasized and the unstressed syllables are much weaker and consequently lose in some cases the full value of their vowels. Unstressed /o/ and /a/ are pronounced as schwa, [ə], e.g. *kómnata* 'room' is pronounced as [kóm̩nətə] and *borodá* 'beard' as [bərədá]. Unstressed /e/ sounds as /I/ in words such as *léca* 'forest (pl)' (Forbes 1956:57-58).

Pashto (Kandahar dialect)

Pashto is another case of lexical accent system with a trochaic character (Revithiadou 1996), as the examples in (8) illustrate. Pashto attracts our interest in this survey because it shortens long syllables and reduces short syllables to schwa in weak foot positions.

(8) *Pashto masculine nouns in -á(e)y* (Penzl 1955)

a. sa(ráy)	b. sa(rí)	'man (direct sg. - direct pl.)'
c. mel(gérey)	d. mel(géri)	'friend (direct sg. - direct pl.)'

The language has a system of seven vowels, /a, a:, e:, o:, ə, i, u/ but, unsurprisingly, long vowels do not occur in weak positions. Penzl (1995:35) distinguishes three types of stress: loud (´), medium (˘) and weak. Long vowels occur only under loud or medium stress, e.g. *zà:ngó:* 'cradle'. Loud stressed /a:, e:, u:/ are in morphophonemic alternation with /a, i, o/, respectively. Moreover, weak stressed /a/ reduces to schwa, [ə]. This reduction is shown in (9), where in the left column stressed full vowels appear, while their unstressed reduced counterparts appear in the right column.

(9) *Reduction of weak syllables*

a. lás	'ten'	yawólás	'eleven'
b. mǎZ	'ram'	mǎZúna	'ram (pl)'
c. kǎ:l	'year'	kǎ:lúna	'year (pl)'

Again, we have shown that unstressed vowels typically reduce and that stressed vowels are typically long.

In the languages analyzed so far, it could be argued that reduction is not of the type usually found in iambic languages. Reduction in iambic languages is more severe, perhaps leading to complete deletion. An example of this is found in his analysis of Eastern Ojibwa. Hayes attributes the severe reductions which happen in the language to its iambic feet. He claims that "severe reduction processes apply to the weak syllables of feet, following the general pattern of increasing durational contrast in an iambic system" (Hayes 1995:216). We will review Hayes's analysis of Eastern Ojibwa and compare it with reduction processes which are equally severe but occur in languages that are trochaic. If reduction in Eastern Ojibwa occurs in order to increase durational contrasts, reductions in trochaic languages are unexpected: in these

languages durational contrasts are disfavored according to the Iambic/Trochaic Law.

Feet are iambs, Hayes argues, and they are assigned from left to right. Furthermore, CVC is light. Examples are given in (13).

- (13) (. *)(. *)(. *)
 ninamadabimi *surface form* nnámádbmi 'we (excl.) sit'
 (. *)(*)(. *)
 oda:we:wigamigw *surface form* dà:wé:gàmik 'a store'

These examples clearly illustrate the reduction of unstressed vowels in Eastern Ojibwa (the final syllable of the last example in (13) is phonetically lengthened, according to Hayes (1995)). In the rest of this section we will present evidence that reduction of unstressed syllables, even in cases where reduction leads to deletion, is not limited to iambic languages, thus weakening the claim laid down in the Iambic/Trochaic Law.

Tiberian Hebrew

Reduction in Tiberian Hebrew is deletion unless an illformed consonant cluster would arise (Rappaport 1984). With respect to reduction Tiberian Hebrew is comparable to Eastern Ojibwa, but Tiberian Hebrew has leftheaded feet, whereas Eastern Ojibwa has righthheaded feet. With respect to the theory advanced in this paper, the reduction of any unstressed vowel is expected, but with respect to the Iambic/Trochaic Law only reduction in righthheaded languages is expected. The fact that reduction in Tiberian Hebrew usually amounts to deletion unless an illformed consonant cluster arises, can point to the fact that the amount of reduction (i.e. reduction or complete deletion) is driven by the phonotactics of a language.

Before reduction in Tiberian Hebrew is analyzed we will present two arguments to support our contention that Tiberian Hebrew is a trochaic language. The first is based on a stress shift of main stress that sometimes occurs. Main stress is on the final syllable if it is closed or if the penultimate syllable is a short open syllable. In other cases main stress is on the penultimate syllable. Only closed syllables are heavy for main stress (McCarthy 1979). If the final two syllables are open and short voweled main stress is retracted to the right.

- (14) *main stress in Tiberian Hebrew*
katába *surface form:* **ka:tbá:**
katábu: *surface form:* **ka:tbú:**

The case *katába* is telling. Stress can end up on the penultimate syllable in two ways: The pattern is either (*katá*)*ba*, with a righthheaded foot, or it is *ka*(*tá*)*ba* with a leftheaded foot. However, the penultimate syllable is an illegitimate stress bearer, therefore, the stress has to move. Assuming that stress stays

within its foot the two outcomes of movement are (*kátø*)*ba* and *ka*(*tøbá*). In (15) stress shift in a trochee and in an iamb are compared.

- (15) *stress shift in trochees and iambs*
- | | | |
|-----------------|-----------|--------------|
| <i>trochees</i> | | <i>iambs</i> |
| (* .) | → (. *) | (. *) |
| CVCV | → CøCV | → CVCø |

Given that *ka*(*tøbá*) is the correct surface form, the conclusion is that main stress in Tiberian Hebrew is assigned in a leftheaded foot. It will also be clear that unstressed vowels delete. The second argument supporting a trochaic analysis of Tiberian Hebrew is epenthesis after gutturals. A guttural must always be followed by a vowel, the quality of which is determined by the vowel preceding the guttural. In this process the second vowel, which is assumed to be a copy of the vowel preceding the guttural, is always reduced. For example *yafímó:d* is realized as *yafĩǎmó:d* (*ǎ* indicates a reduced *a*). Assuming that unstressed vowels are affected by reduction the conclusion is that the domain of this harmony process is a trochaic foot: (*yàfĩǎ*)(*mó:d*).

Having established that the trochee is the foot used in Tiberian Hebrew, we can now turn to the vowel reduction facts. Secondary stress in Tiberian Hebrew is assigned in moraic trochees and both closed syllables and syllables with long vowels are considered to be heavy. In this respect secondary stress differs from main stress, in which only closed syllables are considered to be heavy. Vowels that are not stressed are reduced. Examples are given in (16).

- (16)
- | | |
|--|-------------------------------|
| <i>kə</i> (<i>tobə</i>)(<i>ká</i>) | 'your (masc. sg.) writing' |
| (<i>yik</i>) <i>tə</i> (<i>buu</i>) | 'they (masc.) will write' |
| (<i>way</i>) <i>yə</i> (<i>dab</i>) <i>bə</i> (<i>ruu</i>) | 'and they (masc.) will speak' |

It can be observed that vowels that are unstressed are reduced. This is unexpected under the Iambic/Trochaic Law, but it follows from our generalization in (2).

Proto-Nordic

At some point syncope took place in the trochaic language Proto-Nordic. In this process unstressed vowels have been deleted. The data are given in (17). Riad (1992) analyzed Proto-Nordic as having moraic trochees assigned from left to right. The syncope found in Proto-Nordic shows strong similarities with the reduction process seen above in Eastern Ojibwa. From the point of view of the Iambic/Trochaic Law, the similarity is shocking, however. If syncope in Eastern Ojibwa is intended to enhance durational contrasts, it is likely to suppose that the same holds true for Proto-Nordic. Unfortunately, Proto-Nordic is a trochaic language, which are considered to be unfit to increase durational contrasts according to the Iambic/Trochaic Law.

A few notes on the data. The 'z' developed into an 'r', described by Riad (1992) as 'something like a palatal fricative'; forms marked with an '*' have been reconstructed. All data are taken from Riad (1992).

(14) <i>first syncope period</i>			
*hérðijóoz	>	herðar	'sheperds'
*néθjaz	>	niθiz	'relative'
<i>second syncope period</i>			
sítiz	>	sitr	'sits'
súnu	>	sun	'son (acc.)'

The pattern illustrated in (14) will be clear; unstressed vowels are deleted. If the interpretation of this is that deletion creates a durational contrast, as Hayes does in his analysis of Eastern Ojibwa, this pattern becomes a mystery. Why would a *trochaic* language increase durational contrasts? Under the proposal advanced in this paper, however, the deletion of unstressed vowels is perfectly natural. Unstressed vowels are reduced and stressed vowels are enhanced.

3 Lengthening

Modern Greek

Modern Greek is a *lexical accent system*. In these systems the location of primary accent is the result of a complex interplay of the inherent accentual properties of stems and the diacritic properties of affixes. Stems can be unaccented, accented or they can assign accent to the following morpheme. Suffixes can also be inherently (un)accented or they can remove accent from the domain to which they are attached, assign accent to the preceding syllable, and so on. Such systems have a default pattern, which is found when none of the morphemes of the word asserts its own accentual preference. In Greek the default accent is on the antepenultimate syllable. This is the leftmost syllable of the accentual domain because Greek has a three syllable window. Penultimate and ultimate represent the marked positions of accent. The examples in (15) illustrate how Greek words are accented. The trochaic character of the default pattern, which consists of a foot and a final extrametrical syllable as well as the fact that inherent accent have a trochaic organization, strongly support the idea that the system is purely trochaic (Malikouti-Drachman & Drachman 1988, Revithiadou 1996). It should be mentioned that an essential feature of the system is the accent shifts that words of the default pattern (15a) exhibit when they are combined with accented suffixes (15b).

(15) a. (ánθro)pos	b. an(θrópu)	'man (nom.- gen.sg.)'
c. fan(táros)	d. fan(táru)	'soldier (nom.- gen.sg.)'
e. ura(nós)	f. ura(nú)	'sky (nom.- gen.sg.)'

Accent in Greek is phonetically manifested as stress. Stressed syllables have longer duration and higher amplitude than unstressed syllables, and they are associated with F0 rises (Arvaniti 1991: 52). Arvaniti pursued a number of

experiments which clearly show that stressed syllables are significantly longer than unstressed ones, whether stress is initial or final. We highlight the results of these experiment in the following paragraphs.

Experiment 1 (Arvaniti 1991: 25-34)

In the first experiment Arvaniti was mainly interested in measuring the duration of final vowels in a clash environment. The words of her corpus consisted of two identical CV syllables and were stressed either on the initial or final syllable. There were five combinations all containing one of the five Greek vowels /a, e, i, o, u/. The durational measurements for the pair /papá/ - /pápa/ are given in table 1.

Table 1 : Mean durations (ms) of vowels and syllables of the /papa/ test-word pair. Capitals indicate the stressed syllable of the word.

		vowels	syllable
PA	ms	121	215
pa	ms	92	163
pa	ms	90	154
PA	ms	144	228

Experiment 2 (Arvaniti 1991: 35-47)

Arvaniti's second experiment led to similar results. In this experiment Arvaniti wanted to test, among other things, the duration of syllables in a non clashing context. The findings of the second test for the same pair, namely /pápa/ -/papá/, are presented in table 2.

Table 2 : Mean durations (ms) of vowels and syllables of the /papa/ test-word pair. Capitals indicate the stressed syllable of the word.

		vowels	syllable
PA	ms	130	202
pa	ms	85	133
pa	ms	102	152
PA	ms	131	195

The stressed syllable of /pápa/ is longer than the unstressed one. Vowel durations follow the same pattern as syllable durations for all speakers. Experiment 2 clearly shows that the duration of syllables and vowels increases when they are stressed. The comparison of initial syllables to final ones of the same stress level, shows that initial syllables are either longer or of the same duration as final ones. The longer duration of the stressed syllable of /papá/ in the first experiment, can be attributed to the fact that Greek solves stress clashes

by elongating the first of the clashing syllables and not to lengthening at word final position.

Just as shortening, or reduction, lengthening is unexpected in trochaic systems. According to the Iambic/Trochaic Law lengthening should be limited to iambic systems only. In this section we will present evidence from German, Dutch and Swedish to show that lengthening also occurs in trochaic systems.

The classical analysis of iambic lengthening is the analysis of Hayes of Hixkaryana a Cariban language described by Derbyshire (1985). Closed syllables are heavy. In the vowel inventory there is no distinction between long and short vowels. If a vowel in an open syllable is stressed, it is lengthened, except if it is in final position.

- | | | |
|------|-------------------------------|-----------------------|
| (20) | (khaná:)(nɪh)no | 'I taught you' |
| | (tóh)(kur'ɛ:)(hona | 'to Tohkurye' |
| | (tóh)(kur'ɛ:)(honá:)(hafá:)ka | 'finally to Tohkurye' |

This lengthening of an iambic foot consisting of two light syllables is attributed by Hayes to the Iambic/Trochaic Law. Although the foot (LL) is a well-formed foot, the foot (LH) is even better, since it has a durational contrast. Hence lengthening is understood as the aspiration for the best possible iambic foot. However, this explanation cannot be correct for lengthening is found in trochaic languages as well. Below, lengthening facts from the trochaic languages Dutch, German and Swedish are presented.

Dutch

One of the main conclusions of Nooteboom's (1972) study of vowel duration in Dutch, is that stressed syllables have a longer duration than unstressed syllables. This has been confirmed recently by Tina Langeveld-Cambier. She executed an experiment aimed at the lengthening found in final position. A side effect of this experiment is that it shows that two vowels with the same quality differ only in length. In the word '*rododendron*', the first two vowels are tense, round midvowels; the first one is stressed and the second is not. The second vowel is shorter than the first vowel. This can be explained by assuming that the stressed vowel is lengthened (and the unstressed vowel is somewhat shortened)

Since these vowels are of the same quality, their length difference can only be attributed to their difference in stress. In fact, stress seems to create a durational contrast notwithstanding the trochaic nature of Dutch stress (van der Hulst 1984, Kager 1989).

German

German has been analyzed with a (quantity insensitive) syllabic trochee (Giegerich 1985, Hayes 1986). This means that any two syllables form a foot, regardless of the syllabic make-up of the syllable. Stressed vowels that are [+ATR] are lengthened if they are stressed. The unstressed vowels that are [+ATR] are shortened. This particular lengthening may be regarded as

phonetic and hence not pertinent to the Iambic/Trochaic Law. One of the reasons for assuming that this process is purely phonetic is that there are no other processes referring to vowel length or that there is no phonemic distinction regarding vowel length. However, in Hixkaryana there are no rules referring to vowel length nor is there a phonemic vowel length distinction, yet there is a process of vowel lengthening under stress (Polgárdi 1995). In Hixkaryana this process is used to support the Iambic/Trochaic Law. If Hixkaryana provides support for the Iambic/Trochaic Law then German should count as a counterexample. Our thesis that syllables are lengthened under stress is supported by this example.

Swedish

Swedish provides an example of lengthening in a trochaic language that is phonologically conditioned (Riad 1992:270). At some stage in the history of the Scandinavian languages main stress shifted to the first syllable. At a later stage the first syllable lengthened. Both CVC syllables and the first syllables of CV.CV words lengthened and this can be understood as lengthening under stress. Riad argues that the final consonant is extrametrical, it is ignored for metrical purposes. This would make CVC words too short and lengthening would be the only available solution to save these words from subminimality. This does not hold for the CVCV words, though. For these words, a viable explanation would be to say that the first vowel lengthened under stress and that this has grammaticalized into the language.

(21) *Lengthening in Swedish*

	<i>Old Swedish</i>	<i>Modern Swedish</i>		
		<i>v > vv</i>	<i>c > cc</i>	
a.	nidh	need		'down'
	skip	skeep	skepp	'ship'
b.	ta.la	taa.la		'to speak'
	sku.ta	skoo.ta	skot.ta	'to shovel'

The upshot of this section is that lengthening of stressed syllables also occurs in trochaic languages. The Iambic/Trochaic Law, however, predicts that lengthening only occurs in iambic systems. The durational contrast that is created as a result of lengthening is an iambic property so to say. In Dutch and German lengthening does not affect the grammar; lengthening is a phonetic effect which is found in all stressed syllables. In Swedish lengthening has influenced the grammar of the language. It has been incorporated in the grammar and resulted in the well-formedness condition of Swedish which says that all words must start with a heavy syllable.

Mathimathi

In Mathimathi, an extinct Kulin language with a trochaic stress pattern, stressed syllables are 2 to 2.3 times as long as unstressed syllables (Goedemans 1997). Hayes mentions that stressed syllables must be 1.5 to 2 times as long as unstressed syllable for iambic rhythm to kick in. In

Mathimathi stress can be on the first or on the second syllable (see Gahl 1996, Goedemans 1997 for an account). If the first syllable is stressed and the second is unstressed, the first syllable has an average duration of nearly 120 msec., while the second unstressed syllable is on average a bit longer than 40 msec. This means that the stressed syllable is much longer than twice the unstressed syllable. If the second syllable is stressed and the third is unstressed, then the second syllable is on average 140 msec., while the third syllable is approximately 60 msec. Again the stressed syllable is longer than twice the unstressed syllable. What Mathimathi has in common with languages that have been claimed to exhibit iambic lengthening is that they are pitch accent languages. According to Rice (1992) pitch-accent tends to enhance the lengthening effect of stress.

Although Hayes does not substantiate his claim that stressed syllables in iambic languages should be much longer (i.e. 1.5 to 2 times) than stressed syllables in trochaic languages with (phonetic) evidence from the iambic languages he discusses, we will take this claim seriously and try to account for it. It is known that word final syllables lengthen in general. Suppose that this is a specific manifestation of a general phenomenon which lengthens constituent-final elements. Then there are two forces on constituent final stressed syllables: a) stressed syllables lengthen (see 2 above) and b) constituent-final elements lengthen. The combined effect of these two forces might be responsible for the longer duration of stressed syllables in righthanded languages. In leftheaded feet the picture is different though. The first syllable is lengthened because of its being stressed and the second syllable is shortened because it is unstressed but this shortening is countered by the general lengthening of constituent-final elements. This proposal is pictured in (22).

(22) The joint effects of stress-lengthening and final lengthening

	(σ σ)	(σ σ)
stress-length	↑↓	↓↑
final lengthening	↑	↑

The '↑' means 'make longer', while the '↓' make shorter. In iambs these effects cooperate to make the durational contrasts bigger than in trochees. This explains linguistically why there is a drift towards equal length in trochees and a drift towards uneven length in iambs.

References

- ARVANITI, AMALIA (1991) *The Phonetics of Modern Greek Rhythm and its phonological implications*. Dissertation, University of Cambridge.
- BRAY DE, R.G.A. (1980) *Guide to the Slavonic languages*. Volume 3. Slavic Publishers.
- DAUER, R.M. (1980) The Reduction of unstressed high vowels in Modern Greek. *Journal of the International Phonetic Association* 10, 17-27.
- DERBYSHIRE, DESMOND (1985) *Hixkaryana and Linguistic Typology*. SIL Publications in Linguistics, Summer Institute of Linguistics, Dallas.
- FORBES, NEVILL (1956) *Russian Grammar*. Oxford, Clarendon Press.

- GIEGERICH, HEINZ (1985) *Metrical Phonology and Phonological Structure*. Cambridge Studies in Linguistics 43, Cambridge University Press, Cambridge.
- GOEDEMANS, R. (1997) Putting the Mathimathi stress rule in its proper perspective. Ms. Leiden University, to appear in *Australian Journal of Linguistics*.
- HAYES, BRUCE (1995) *Metrical Stress Theory: Principles and Case Studies*. University of Chicago Press, Chicago.
- HULST, HARRY VAN DER (1984) *Syllable Structure and Stress in Dutch*. Foris, Dordrecht.
- KAGER, RENÉ (1989) *A Metrical Theory of Stress and Destressing in English and Dutch*. Foris, Dordrecht.
- KAGER, RENÉ (1993) Alternatives to the Iambic/Trochaic Law. *Natural Language and Linguistic Theory*, 11, 381-432.
- MALIKOUTI-DRACHMAN, ANGELIKA, AND GABARELL DRACHMAN (1988) Stress in Greek (Tonismos sta Ellinika). *Studies in Greek Linguistics* 9, 127-43.
- MARGARITI-ROGA (1985) *Phonological Analysis of the Siatista dialect (Phonologiki Analsi tu Siatistinu Idiomatos)*. Dissertation, University of Thessaloniki.
- MAYO, PETER (1976) *A Grammar of Byelorussian*. The Anglo-Byelorussian Society in Association with the Department of Russian and Slavonic Studies, Sheffield, University of Sheffield.
- MCCARTHY, JOHN (1979) *Formal Problem in Semitic Phonology and Morphology*. Doctoral Dissertation, MIT. [published by Garland Press, New York, 1985]
- NOOTEBOOM, S.G. (1972) *Production and Perception of Vowel Duration. A Study of the Durational Properties of Vowels in Dutch*. Dissertation, University of Utrecht.
- PENZL, H. (1995) *A Grammar of Pashto; A Descriptive Study of the Dialect of Khandar, Afganistan*. American Council of Learned Societies, Washington D.C..
- REVITHIADOU, ANTHI (1996) Marked feet in the Pool. Ms. Rijksuniversiteit Leiden/HIL. To appear in *Proceedings of HILP 3*.
- RIAD, TOMAS (1992) *Structures in Germanic Prosody: A Diachronic Study with Special Reference to the Nordic Languages*.
- RICE, CURTIS (1992) *Binarity and Ternarity in Metrical Theory: Parametric Extensions*. Dissertation. University of Texas at Austin.
- THEOPHANOPOULOU-KONTOU, D. (1973) *Elements of Phonology and Phonetics (Stoixeia Phonitikis-Phonologias)*. Athens.
- VIJVER, RUBEN VAN DE (1996) The accidental iamb. In Nespor and Smith (eds.), *Dam Phonology*. HIL Publications 3, HAG, Den Haag.
- WIESE, RICHARD (1996) *The Phonology of German*. Oxford University Press, Oxford.

Subjacency Forever

Norvin Richards
MIT

1. Problems: levels and taxes

Baker (1970) notes that sentences of the following form are ambiguous in English:

- (1) [Who knows [who bought what]]?

Here *what* can have scope either with the *who* in the embedded clause or with the *who* in the matrix clause. The availability of the latter reading is somewhat surprising in light of the ill-formedness of (2):

- (2) *[What do you know [who bought t]]?

That is, the process whereby *what* gets its scope in (1) is apparently immune to the constraint which rules out (2). One can imagine a number of ways in which the contrast between (2) and the wide-scope reading of (1) might be explained; I will concentrate here on two.

The first, which has been fairly well accepted in much of the literature on this problem, would be to say that the LF movement whereby *what* gets its scope in (1) is immune to Subjacency. Let us refer to this as the “levels approach”. One argument for this approach comes from Huang (1982), who notes that in Chinese, a language in which wh-movement is apparently not overt, no wh-island effects are observed. The same is true in Tibetan ((3a) from Huang 1982, 267; (3b) from Ngawang Jorden, p.c.)¹:

- (3) a. Ni xiang-zhidao shei mai-le sheme?
you wonder who bought what
‘What do you wonder who bought?’
b. Khong-gi khyedrang-la [su -s gare gzigs-pa]
he ERG you DAT who ERG what buy that
bka’’dri- gnang- pa- red?
question do-HON PAST AGR
‘What did he ask you who bought?’

Another possible way of accounting for the distinction between (1) and (2), argued for in Brody 1995b and Richards 1996, would be to say that Subjacency need only be obeyed once per wh-comp; that is, in (1), the local movement of *who* to the matrix Comp satisfies Subjacency, rendering all subsequent wh-movements to that site (for instance, the movement of *what*) immune to Subjacency. I will refer to this as the “Subjacency Tax approach”. Evidence for this approach comes from the contrasts in (4) (Japanese, from Watanabe 1992) and (5) (Bulgarian, from Roumyana Izvorski and Roumyana Slabakova, p.c.):

- (4) a. ?? John-wa [Mary -ga nani -o katta ka dooka] siritagatte-iru no?
 John TOP Mary NOM what ACC bought whether know-want Q
 'What does John want to know whether Mary bought?'
 b. John-wa [Mary-ga nani -o katta ka dooka] dare -ni tazuneta no?
 John TOP Mary NOM what ACC bought whether who DAT asked Q
 'Who did John ask whether Mary bought what?'
- (5) a. * Koj-a kniga; otrece senatorat [malvata če iska da zabrani ti]?
 which book denied the-senator the-rumor that wanted to ban
 'Which book did the senator deny the rumor that he wanted to ban?'
 b. ? Koj senator koja kniga; otrece [malvata če iska da zabrani ti]?
 which senator which book denied the-rumor that wanted to ban
 'Which senator denied the rumor that he wanted to ban which book?'

(4a) and (5a) are both Subjacency violations. In the (b) sentences we can see that adding an additional wh-word outside the Subjacency island improves the structure in both cases. These are both languages in which all wh-movement is apparently done on a single syntactic level, so the approach outlined above, in which the relevant factor is whether the movement is overt or covert, has nothing straightforward to say about these cases.²

Neither account deals easily with the contrast between Japanese and Chinese, shown again in (6):

- (6) a. ??John-wa [Mary -ga nani -o katta ka dooka] siritagatte-iru no?
 John TOP Mary NOM what ACC bought whether know-want Q
 'What does John want to know whether Mary bought?'
 b. Ni xiang-zhidao shei mai-le sheme?
 you wonder who bought what
 'What do you wonder who bought?'

Languages which do all their wh-movement covertly seem to be divisible into two classes; a class with wh-island effects (including Japanese and Korean) and a class without them (including Chinese and Tibetan). What property accounts for this distinction?

Rudin (1988) notes that languages which do all their wh-movement overtly are also divisible into a class which exhibits wh-island effects (including Serbo-Croatian and Polish) and a class which lacks them (including Bulgarian and Rumanian):

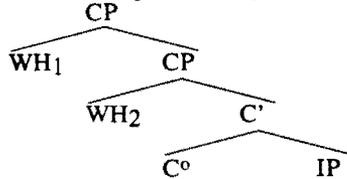
- (7) (Serbo-Croatian, from Rudin 1988, 459)
 a. * Šta si me pitao ko može da uradi?
 what AUX me asked who can to do
 'What have you asked me who can do?'
 (Bulgarian, from Rudin 1988, 457)
 b. ? Koj-a ot tezi knigi se čudiš koj znae koj prodava?
 which of these books wonder-2s who knows who sells
 'Which of these books do you wonder who knows who sells?'

The hypothesis which will be pursued in this paper will be that the difference between the Japanese class and the Chinese class is the same as the difference between the Serbo-Croatian class and the Bulgarian class. We will see that there are independent reasons for the Chinese class of languages to lack wh-islands, and the levels approach is therefore undermined.

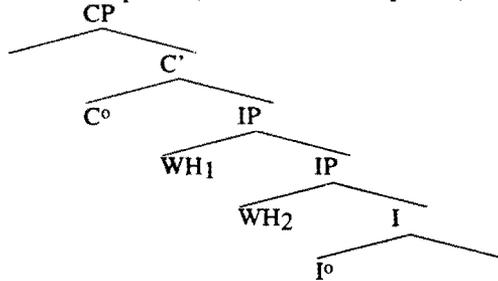
2. CP-Absorption and IP-Absorption

Adapting somewhat the proposal of Rudin (1988), I will hypothesize that multiple wh-movement can take place either by movement to multiple specifiers of CP³, as in (8a), or by movement to multiple specifiers of an IP projection, as in (8b).

- (8) a. “CP-absorption” (Bulgarian, Chinese)



- b. “IP-absorption” (Serbo-Croatian, Japanese)



I will refer to the former type as “CP-absorption languages”, and to the latter as “IP-absorption languages”. Bulgarian and Chinese are CP-absorption languages; Serbo-Croatian and Japanese are IP-absorption languages.

CP-absorption languages have more or less familiar properties; wh-movement is always to a specifier of CP, and is always A-bar movement. IP-absorption languages, on the other hand, have somewhat more exotic properties. Here wh-movement most closely resembles the scrambling found in languages like Hindi (cf. Mahajan 1992) and Japanese (cf. Saito 1992); local wh-movement has certain properties of A-movement, while long-distance wh-movement uniformly acts like A-bar movement, presumably because A-chains are subject to stricter locality principles. In some IP-absorption languages, a single wh-word apparently moves

obligatorily to Spec CP (Serbo-Croatian, for example, appears to be such a language, although Hungarian is not). I will not speculate here on the force driving this move.

In the next section I will discuss the differences between IP-absorption languages and CP-absorption languages, and try to show that Japanese and Chinese do indeed differ in this regard.

3. Diagnostics for CP- and IP-absorption

3.1 Wh-islands

First, let us consider how the posited structures for CP- and IP-absorption languages account for the distribution of wh-island effects.

We have seen that some IP-absorption languages make use of Spec CP as a landing site for wh-movement; Serbo-Croatian apparently requires one wh-word to raise to Spec CP. Now we are in a position to give an account of wh-islands, essentially following Rudin (1988), and Comorovski (1986). Suppose that wh-movement past a filled Spec CP is universally barred, for familiar reasons having to do with considerations of Shortest Move. The only languages which will allow wh-movement out of a question, then, will be ones in which CP can have multiple specifiers, so that wh-movement need never skip a CP projection; there will always be a specifier of CP available as an escape hatch. In IP-absorption languages, on the other hand, it is IP which has multiple specifiers, and CP has only one. A single wh-word which has been forced to move to Spec CP, then, blocks further wh-movement past that specifier position. Thus, IP-absorption languages should exhibit wh-island effects whenever a single element occupies Spec CP, while CP-absorption languages should lack such effects.

3.2 Scrambling

The IP-absorption languages all exhibit a form of local scrambling which fixes weak crossover violations:

(Serbo-Croatian, from Milan Mihaljevic)

- (9) a. ??Njegovi susjedi ne vjeruju nijednom politicaru;
 his neighbors not trust no politician
 'His neighbors trust no politician'
 b. Nijednom politicaru; njegovj susjedi ne vjeruju tj
 no politician his neighbors not trust

(Japanese, from Saito 1992, 73)

- (10)a. ?* Soituj-no hahaoya-ga darej-o aisiteiru no?
 guy GEN mother NOM who ACC loves Q
 'Who does his mother love?'
 b. ? Darej-o soituj-no hahaoya-ga aisiteiru no?
 who ACC guy GEN mother NOM loves Q

(Hungarian, from Kiss 1994, 22)

- (11)a.* Nem szeret az proj anyja mindenkit;
 not loves the mother-his everybody-ACC
 'His mother does not love everybody'
 b. Nem szeret mindenkit; az proj anyja
 not loves everybody-ACC the mother-his

CP-absorption languages, on the other hand, apparently lack such a form of scrambling; scrambling is either absent entirely or is A-bar movement:

(Bulgarian, from Roumyana Slabakova)

- (12) a. *Majka mu obicha vseki chovek
 mother his love every person
 'His; mother loves everyone;'
 b. *Vseki chovek obicha majka mu
 every person love mother his

(Chinese, from Hooi-Ling Soh)

- (13)a. * Tade mama ai meigeren
 his mother love everyone
 'His; mother loves everyone;'
 b. * Meigeren tade mama ai
 everyone his mother love

The pattern, then, seems to be that all and only languages which allow local A-scrambling are IP-absorption languages; this is true regardless of the level on which wh-movement occurs. On the assumption that both IP-absorption and local A-scrambling involve either adjunction to IP or movement into multiple specifiers of IP, this result has a certain intuitive appeal; if a language allows this kind of movement, it uses it both for scrambling and for wh-movement, and if not, neither scrambling nor IP-absorption will be found.

3.3. Superiority

Rudin (1988) notes that in languages like Bulgarian (CP-absorption languages, in this theory), the ordering of fronted *wh*-phrases is subject to a rigid ordering, which she attributes to Superiority:

- (14)a. Koj kogo e vidjal?
 who whom AUX seen
 'Who saw whom?'
 b. *Kogo koj e vidjal?
 (15)a. Koj kude udari Ivan?
 who where hit Ivan
 'Who hit Ivan where?'
 b. *Kude koj udari Ivan?
 where who hit Ivan

In an IP-absorption language like Serbo-Croatian, on the other hand, Superiority effects obtain only for long-distance movement, and not for local movement, as Boskovic (1995) points out:

- (16)a. Ko je koga vidjeo?
 who AUX whom seen
 'Who saw whom?'
 b. Koga je ko vidjeo?
 (17)a. Ko si koga tvrdio da je istukao?
 who AUX whom claimed that AUX beaten
 'Who did you claim beat whom?'
 b. *Koga si ko tvrdio da je istukao?
 whom AUX who claimed that AUX beaten

The sense in which these restrictions on ordering may be attributed to Superiority is not a straightforward one, but I will not discuss the matter here, for reasons of space (cf. Richards (to appear) for some further discussion). For our purposes it is sufficient to note that the differences between Bulgarian and Serbo-Croatian are accounted for by the theory developed here, assuming that Superiority constrains A-bar movement but not A-movement. All Bulgarian *wh*-movements are A-bar movements, being movements to Spec CP; in Serbo-Croatian, on the other hand, *wh*-movement is adjunction to an IP-level projection, and may be an A-movement if it is sufficiently local. The lack of Superiority effects for local movement in Serbo-Croatian (that is, in IP-absorption languages) therefore follows.

Interestingly, a similar asymmetry between local and long-distance movement seems to be present in the LF-moving IP-absorption languages. Japanese Anti-superiority, like the Superiority effects in Serbo-Croatian, is stronger (for some speakers) with long-distance movement than it is with local movement (Minoru Fukuda, Shigeru Miyagawa, p.c.):

- (18)a. John -ga nani -o naze katta no?
John NOM what ACC why bought Q
'What did John buy why?'
b. ? John-ga naze nani -o katta no?
John NOM why what ACC bought Q
- (19)a. Mary -ga [John -ga nani -o naze katta to] omotteiru no?
Mary NOM John NOM what ACC why bought that thinks Q
'What did Mary think John bought why?'
b. * Mary-ga [John -ga naze nani -o katta to] omotteiru no?
Mary NOM John NOM why what ACC bought that thinks Q

Thus, the Serbo-Croatian and Japanese equivalents of Superiority seem to behave similarly, in that they constrain only long-distance movement; according to the story developed here, this is because only long-distance movement has A'-properties in these languages. The prediction of this account would be that Chinese Superiority, like Bulgarian Superiority, would be equally strong locally and long-distance. Chinese word order is too rigid to test this; no alternatives parallel to those in (18-19) can be constructed. On the other hand, in Tibetan, another LF-moving CP-absorption language, scrambling is possible, and we find strong local Superiority effects (Ngawang Jorden, p.c.):

- (20)a. Bkrashis-lags -gi gyag garebyadnas gzigs-gnang-pa -red?
Tashi HON ERG yak why buy -HON -PAST -AGR
'Why did Tashi buy a yak?'
b. Bkrashis-lags -gi garebyadnas gyag gzigs-gnang-pa -red
Tashi HON ERG why yak buy HON PAST AGR
- (21)a. Bkrashis-lags -gi gagi garebyadnas gzigs-gnang-pa -red
Tashi HON ERG which why buy HON PAST AGR
'Why did Tashi buy what?'
b. * Bkrashis-lags -gi garebyadnas gagi gzigs-gnang-pa -red
Tashi HON ERG why which buy HON PAST AGR

(20a-b) show that scrambling of *garebyadnas* 'why' over the direct object is possible in principle, but (21a-b) show that it is impossible if the direct object is itself a wh-word. Thus, the Tibetan equivalent of Japanese Anti-superiority strongly constrains local movement, as we expect on the hypothesis that Tibetan is like Bulgarian in that all wh-movement is A-bar movement to a Spec CP position.

3.4. Weak Crossover

Another asymmetry between local and long-distance movement in IP-absorption languages appears in the domain of weak crossover. CP-absorption languages like Bulgarian have weak crossover effects of a fairly familiar kind (Roumyana Slabakova, p.c.):

- (22)a. Koj obicha majka si?
 who loves mother his
 'Whoj loves hisj mother?'
- b. *Kogo obicha majka mu?
 who loves mother his
 'Whoj does hisj mother love?'

In IP-absorption languages like Hungarian and Serbo-Croatian, on the other hand, weak crossover effects are found only long-distance, not locally: (Hungarian from Kiss 1987, 208, and Brody 1996; Serbo-Croatian from Snjezana Kordic):

- (23)a. Ki szereti az anyjat?
 who loves the mother-his-ACC
 'Whoj loves hisj mother?'
- b. Kit szeret az anyja?
 who-ACC loves the mother-his
 'Whoj does hisj mother love?'
- c. * Kit gondol az anyja hogy Mari szeret?
 who-ACC thinks the mother-his that Mary loves
 'Whoj does hisj mother think that Mary loves?'
- (24)a. Tko voli svoju majku?
 who loves his-ACC mother-ACC
 'Whoj loves hisj mother?'
- b. Koga voli njegovu majku?
 who loves his-NOM mother-NOM
 'Whoj does hisj mother love?'
- c. * Koga njegovu majku misli da Marija voli?
 who his-NOM mother-NOM thinks that Maria loves
 'Whoj does hisj mother think that Mary loves?'

A surprising fact, given the theory developed here, is the presence of weak crossover effects in IP-absorption languages like Japanese (Saito 1992, 73):

- (25) ?* Soituj-no hahaoya-ga darej -o aisiteru no?
 guy GEN mother NOM who ACC love Q
 'Who does his mother love?'

It is not clear why *dare* cannot adjoin to IP at LF in a position higher than the pronominal variable it binds, thus obviating the weak crossover violation. One possible answer will be outlined in the next section.

3.5. Wh-movement and QR

IP-absorption, as developed here, is syntactically reminiscent of QR, in that it involves multiple adjunction to IP in order to establish scope relations. In some

languages in which IP-adjunction occurs overtly, movement does indeed seem to have effects on scope relations:

(Hungarian, from Kiss 1994, 71)

- (26)a. Mindenki két lányt is meg táncoltatott
 everyone two girl-ACC even PREV danced
 'Everybody danced with two (potentially different) girls'
 b. Két lányt is mindenki meg táncoltatott
 two girl-ACC even everyone PREV danced
 'Two girls (the same two girls) were danced with by everybody'

(Japanese, adapted from Kuroda 1971)

- (27)a. Dareka -ga daremo -o aisiteru
 someone NOM everyone ACC loves
 'Someone loves everyone' ($\exists >> \forall, * \forall >> \exists$)
 b. Daremo_i -o dareka -ga t_j aisiteru
 everyone ACC someone NOM loves
 'Someone loves everyone' ($\exists >> \forall, \forall >> \exists$)

Furthermore, IP-adjoined wh-words in Hungarian demonstrably occupy a position which is also used as a landing site by a certain class of quantificational elements; such quantifiers have their scopes determined by movement to this position (see Kiss 1987, 1994 for discussion). In (28) we see that both wh-movement and this form of overt QR trigger inversion of the verb with a preverb, a standard test for occupying this position (from Kiss 1994 (37, 64)):

- (28)a. Ki hívta fel Jánost?
 who called PREV John-ACC
 'Who called up John?'
 b. János kevés fogást kóstolt meg
 John few dish-ACC tasted PREV
 'John tasted few dishes'

Japanese and Chinese are both "rigid scope" languages; the scope of quantifiers is apparently entirely determined by their surface position, so that (29a-b) are both unambiguous, with the subject QP taking scope over the object QP.

(Chinese & Japanese, from Aoun and Li 1993, 365)

- (29)a. (Yaoshi) yige ren piping meigeren...
 if one man criticize everyone
 '(If) someone criticized everyone...'
 b. Dareka -ga daremo -o semeta
 someone NOM everyone ACC criticized
 'Someone criticized everyone'

According to the theory developed here, Chinese and Japanese differ in that Japanese uses the same syntactic mechanism, IP-adjunction, to assign scope to

quantifiers and to *wh*-words, while Chinese uses two different syntactic mechanisms: IP-adjunction and substitution to Spec CP. Interestingly, the “rigid scope” property of quantifiers is extended to *wh*-words in Japanese, but not in Chinese, as Aoun and Li (1993) point out; the ill-formedness of (29b) may be attributed to the inability of the *wh*-word to take scope over the other operators in the sentence:

- (30) a. Meigeren dou maile shenme?
 everyone all bought what
 ‘What did everyone buy?’
 b. *Daremo -ga nani -o kaimasita ka?
 everyone NOM what ACC bought Q
 ‘What did everyone buy?’

Although I have no account of “rigid scope” to offer, it seems clear that the theory developed here makes the difference between Chinese and Japanese look less surprising; the generalization, apparently, is that LF IP-adjunction in these languages cannot result in a change of scope relations.

Rigid scope might also be responsible for the ill-formedness of (25), repeated as (31):

- (31) ?* Soituj-no hahaoya-ga darej -o aisiteru no?
 guy GEN mother NOM who ACC love Q
 ‘Who does his mother love?’

Whatever our eventual account of scope rigidity might be, we might expect it to say that *dare* is unable to bind any variables at LF which it cannot bind in the overt syntax; the surprising ill-formedness of (31) would then follow.

4. Conclusion

In this paper I have tried to show that Rudin’s (1988) observation that multiple *wh*-movement languages can be classified in two types (here referred to as “CP-Absorption” and “IP-Absorption” languages) holds for languages which do *wh*-movement covertly as well. Briefly, the claim here has been that languages like Bulgarian and Chinese perform movement to multiple specifiers of CP, while languages like Serbo-Croatian and Japanese perform *wh*-movement by multiple adjunction of IP, an operation which is syntactically similar both to Japanese scrambling and QR. Data from a variety of areas (including not only *wh*-island effects but also Superiority effects, interactions between *wh*-words and quantifiers, and the availability of local A-scrambling) seem to lend support to this claim. The empirical observations are summarized in the table below:

	Japanese	Serbo-Croatian	Hungarian	Chinese	Tibetan	Bulgarian
wh-islands	yes	yes	yes	no	no	no
local A-scrambling	yes	yes	yes	no		no
local Superiority	no	no	no		yes	yes
local WCO	yes	no	no	yes		yes
wh-mvmt=QR	yes		yes	no		

To the extent that the above analysis is well-founded, it undermines the claim that Subjacency does not constrain LF movement. I have argued here that the strongest single piece of evidence for this claim, the absence of Subjacency effects in Chinese, can and should be derived in a different way. This approach also seems inconsistent with accounts in which there is no LF movement (e.g., Reinhart 1993, 1995, Tsai 1995); to the extent that overt and covert movement can be shown to have similar properties, an account which assumes radically different syntactic mechanisms for dealing with moved wh and wh-in-situ seems undesirable.

Acknowledgements

Many thanks to the following people for their help with the facts of their languages: Roumyana Izvorski, Ani Petkova, Virginia Savova, and Roumyana Slabakova (Bulgarian), Hooi-Ling Soh and Wei-Tien Dylan Tsai (Chinese), Takako Aikawa, Minoru Fukuda, and Shigeru Miyagawa (Japanese), Michael Brody, Julia Horvath, and Anna Szabolcsi (Hungarian), Snjezana Kordic and Milan Mihaljevic (Serbo-Croatian), and Ngawang Jorden (Tibetan). For more general comments on the paper I would also like to thank Yoo-Kyung Baek, Noam Chomsky, Alec Marantz, Martha McGinnis, David Pesetsky, Shigeru Miyagawa, and audiences at MIT, CONSOLE V, and WECOL '96. Any faults that I have managed to sneak past all these people are purely my own responsibility. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship.

¹ Both of these examples, irrelevantly for our purposes, also have a reading in which both wh-words have embedded scope.

² One might, as Watanabe (1992) does, postulate a class of movements which are overt but invisible to deal with the Japanese facts in (4); on such an account, Subjacency applies to wh-movement in (4a) because *nani* 'what' actually moves overtly, though invisibly. (4b) would then be well-formed because the wh-word which moves overtly is *dare*; *nani* moves covertly in this case (and therefore invisibly). To deal with the facts in (5), one could presumably make a similar claim in reverse, postulating a class of movements which are visible but covert. This would be the kind of movement which applies to *koja kniga* 'which book' in (5b); Subjacency does not apply here because the movement is covert (although it can be seen). Maneuvers of this kind will allow us to maintain the generalization that all and only overt movements are subject to Subjacency, but carry a certain risk of rendering that generalization vacuous.

³ Alternatively, this movement might involve multiple adjunction to CP; I have no evidence bearing on this distinction, if indeed such a distinction exists.

References

- Aoun, J., & Y.-H. A. Li. 1993. On some differences between Chinese and Japanese wh-elements. *Linguistic Inquiry* 24, 365-72.
- Baker, C. 1970. Notes on the description of English questions. *Foundations of Language* 6, 197-219.
- Beck, S. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4, 1-56.
- Bošković, Z. 1995. Superiority effects with multiple wh-fronting in Serbo-Croatian. Ms., University of Connecticut.
- Brody, M. 1995a. Focus in Hungarian and bare checking theory. *Tübingen Working Papers in Linguistics*.
- . 1995b. *Lexico-logical form: a radically minimalist theory*. MIT Press, Cambridge.
- Chomsky, N. 1995. *The minimalist program*. MIT Press, Cambridge.
- Comorovski, I. 1986. Multiple wh-movement in Romanian. *Linguistic Inquiry* 17, 171-177.
- Horvath, J. 1986. *FOCUS in the theory of grammar and the syntax of Hungarian*. Foris, Dordrecht.
- Huang, C.-T. J. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.
- Kiss, K. E. 1987. *Configurationality in Hungarian*. Reidel, Dordrecht.
- . 1994. Sentence structure and word order. *Syntax and semantics 27: The syntactic structure of Hungarian*, ed. by F. Kiefer & K. E. Kiss, 1-90. Academic Press, San Diego.
- Kuroda, S.-Y. 1971. Remarks on the notion of subject with reference to words like also, even, or only, part II. *Annual Bulletin of the Logopedics and Phoniatrics Research Institute* 4, 127-152.
- Mahajan, A. 1992. The A/A-bar distinction and movement theory. Doctoral dissertation, MIT.
- Marác, L. 1989. Asymmetries in Hungarian. Doctoral dissertation, Rijksuniversiteit Groningen.
- Reinhart, T. 1993. Wh-in-situ in the framework of the Minimalist Program. *OTS Working Papers*, Research Institute for Language and Speech, Utrecht University, Utrecht, the Netherlands.
- . 1995. Interface strategies. *OTS Working Papers*, Research Institute for Language and Speech, Utrecht University, Utrecht, the Netherlands.
- Richards, N. 1996. The Principle of Minimal Compliance. *Proceedings of CONSOLE 4*, ed. by J. Costa, R. Goedemans, & R. van de Vijver, 265-282.
- . To appear. Shortest Moves to (Anti-)Superiority. To appear in *Proceedings of WCCFL XVI*.

- Rudin, C. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6, 445-501.
- Saito, M. 1982. Scrambling, topicalization, and strong crossover. Ms., MIT.
- . 1992. Long-distance scrambling in Japanese. *Journal of East Asian Linguistics* 1, 69-118.
- Sauerland, U. 1996. The interpretability of scrambling. Ms., MIT.
- Takahashi, D. 1993. Movement of wh-phrases in Japanese. *Natural Language and Linguistic Theory* 11, 655-678.
- Tsai, W.-T. D. 1995. On economizing the theory of A-bar dependencies. Doctoral dissertation, MIT.
- Watanabe, A. 1992. Subjacency and S-structure movement of wh-in-situ. *Journal of East Asian Linguistics* 1, 255-291.

**Evidence for the Optional Tense Hypothesis:
tense in subordinate clauses in the acquisition of English***

Taylor Roberts
Massachusetts Institute of Technology

Introduction

When children acquiring English begin to form sentences at around two years old, they produce both inflected and uninflected main clauses for several months, despite the fact that the adult grammar requires that main clauses be inflected. Examples of such **optional infinitives** (OIs) are given in (1):

- | | | | |
|-----|----|-------------------------|----------------|
| (1) | a. | Adam drop it train | (Adam 2;5.12) |
| | b. | Write dat piece a paper | (Adam 2;6.17) |
| | c. | Hold baby | (Sarah 2;4.26) |
| | d. | He bite me | (Sarah 2;6.13) |

Two proposals have recently been put forth to explain this phenomenon. The **Truncation Hypothesis** of Rizzi (1994*a*, 1994*b*) and others suggests that children may produce clauses exhaustively dominated not by a CP (which is the root node commonly assumed to represent adult clauses), but rather by a lower node. In contrast, the **Optional Tense Hypothesis** of Wexler (1994) allows the child to omit only the TP projection while retaining higher functional projections. Each hypothesis accounts for a set of generalizations associated with the OI stage; for example, the cooccurrence of null subjects, *wh*-movement, negation, infinitival *to*, and so on. (For some recent discussion, see several of the papers in Schütze et al. (1995).) Furthermore, these hypotheses predict properties of subordinate clauses that seem not to have been discussed or tested, and so this paper undertakes this area of inquiry. The speech of two children is examined in order to determine the extent of correct tense marking in subordinate clauses. It will be shown that the Optional Tense Hypothesis yields more accurate predictions about tense marking in the earliest subordinate clauses that appear in children's speech.

Adult clauses are generally assumed to have a structure similar to (2). See, for example, Chomsky (1993):

- (2) [CP . . . [AgrSP . . . [TP . . . [AgrOP . . . [VP . . .

Of particular relevance to the OI phenomenon is the medial projection TP (Tense Phrase), since tense is crucially missing from OIs. Adult English shows complementarity between its tense morphemes: a clause may be finite—in which case T bears a [\pm past] feature—or nonfinite, in which case T is empty (as in

small clauses) or headed by the infinitival particle *to*. In order to account for the absence of tense in OIs, the Truncation Theory allows (2) to be abbreviated, by permitting the child to represent a sentence with a root node lower than CP; since tense is absent, the root node for an OI would be AgrOP or VP. Sentence (1d) would thus have the following S-structure representation:

(3) [vp he [v' bite me]]

Because TP has been truncated from (3), there is no T node bearing [\pm past] features to which the verb may raise and be inflected, as it would be in the adult sentence *he bit me* or *he is biting me*. Note also that the absence of the infinitival *to* particle is also expected, since there is no T node under which it may appear. The specific manner in which the child's grammar differs from that of the adult is that the following principle is not yet operative in the child's grammar:

(4) Root = CP

Until this principle appears in the grammar (whether by acquisition or by maturation), the child is free to take categories other than CP for the root node. Rizzi (1994a: 162) and others assume that the principle in (4) expresses the fact that speech generally takes the form of propositions, and that the canonical structural realization of the proposition is CP.

A desirable prediction of the Truncation Hypothesis is that it explains the rarity of the cooccurrence of null subjects with *wh*-movement; in other words, why the null subject sentence in (5b) seldom appears as a variant of (5a). On the other hand, *in situ wh*-words frequently appear with null subjects, as in (5c).

- (5) a. Where Daddy go? (Adam 2;3.18)
 b. Where go?
 c. See what bear? (Adam 2;4.3)

Since null subjects are otherwise generally permitted in early English, sentences like (5b) should be as common as are null subjects in non-*wh* sentences, and yet it is reported that they are not. The failure of *wh*-movement to correlate with null subjects receives a straightforward explanation under the Truncation Hypothesis. Rizzi (1994a, 1994b) suggests that null subjects are allowed in the acquisition of such non-null subject languages as English and French because the higher levels of structure have been truncated, and so a null subject in a lower projection, such as Spec/IP or Spec/AgrSP, may be identified by a referent from the discourse; if there were projections above IP, the null subject would seek a c-commanding controller within its own sentence, and a failure to

find one would result in ungrammaticality, since the null subject would not be bound by any referent. Because *wh*-movement entails the presence of CP (the *wh*-word moves to Spec/CP), a null subject would not be licit in Spec/IP, since it would fail to find an antecedent in its own clause. A rarely attested sentence like (5b) would have an S-structure like the following (ignoring irrelevant details):

(6) [CP where_i [IP e [VP go t_i]]]

As the null subject does not have an antecedent within its own clause in this structure, it is ill formed, and hence such a structure is rarely produced in children's speech. When the *wh*-word remains *in situ*, however, as it does in (5c), there is no overt evidence for the presence of the CP projection, and so this sentence would be assigned the following truncated structure:

(7) [IP e [VP see what bear]]

There being no structure above the null subject in Spec/IP, it is free to find its antecedent in the discourse, and hence this structure is well formed for (and frequently produced by) the child. A paradigm similar to that in (5) may be constructed also for verb-second languages like Dutch and German; null subjects seldom appear in verb-second sentences, since the verb has presumably moved to C0, making it impossible for a null subject below, in Spec/IP, to be identified within its own clause.

Note, incidentally, that Rizzi's (1994a, 1994b) analysis is characteristic of what Wexler (1994: 305) calls the "Strong Universal Grammar" view of language acquisition: children are assumed to know most of the grammatical principles of the language, although initially it may be difficult for the linguist to discern this knowledge; because sentences that are ungrammatical (from the perspective of the adult grammar) result from the small number of principles that the children do not yet control, the underlying grammatical knowledge that the children do possess is obscured, and frequently this knowledge may be inferred only by observing subtle asymmetries in production as between such structures as those exemplified in (5).

In contrast to the Truncation Hypothesis representation shown in (3), the Optional Tense Hypothesis offers the following, fuller representation of (1d):

(8) [CP [AgrSP [AgrOP [VP he [V' bite me]]]]]

As in the Truncation Hypothesis representation (3), the Optional Tense Hypothesis does not offer a TP projection in which tense elements or infinitival *to* may be borne. With respect to the other functional projections, however, the

child has full competence—a highly desirable characteristic for both empirical and theoretical reasons, as discussed by Wexler (1994), Poeppel and Wexler (1993), and the references there. Both hypotheses allow T to be absent in early grammatical structure, but the absence of T in the Truncation Hypothesis entails the absence of several higher projections, while the absence of T in the Optional Tense Hypothesis is taken as an axiom. Both theories predict the absence of tense in main clauses, as well as the cooccurrence (or not) of other properties in OI sentences, but the strongest version of the Truncation Hypothesis makes the additional prediction that only the tense of main clauses may be omitted; subordinate clauses—because their maximal projections (CPs) are selected by a predicate in the matrix clause—should show adult-like tense properties. In contrast, the Optional Tense Hypothesis suggests that TP may be omitted more broadly, predicting that subordinate clauses in children's speech will display the same degree of incorrect inflection in embedded clauses during the OI stage as do main clauses.

Method

Subjects

The correctness of tense marking in subordinate clauses is examined in the speech of two children, Adam and Sarah, both of whom speak American English. The data originate from computerized transcripts of the Brown (1973) corpora, which form part of the CHILDES database (MacWhinney and Snow, 1985, 1990, 1992; MacWhinney, 1995). Adam's speech is examined from 2;3.4 through 2;11.0, and Sarah's from 2;3.5 through 2;10.11.

Procedure

The speech of each child was examined starting from the earliest appearance of subordinate clauses, while—crucially—each child is still clearly in the OI stage with respect to main clauses. Subordinate clauses in Adam's speech appear in the earliest records available for him (around two years and three months):

- | | | | |
|-----|----|------------------------------|---------------|
| (9) | a. | Go get it | (Adam 2;3.4) |
| | b. | Want sit down | (Adam 2;3.4) |
| | c. | Go belong | (Adam 2;3.18) |
| | d. | Adam put boot on looking for | (Adam 2;5.12) |
| | e. | Laughing being cowboy | (Adam 2;5.12) |
| | f. | Look birdie fly | (Adam 2;5.12) |
| | g. | Gon (t)a ride dat | (Adam 2;6.3) |

As may be seen from the above examples, it is often hard to determine the structure of the subordinate clause. For example, *Go get it*, while grammatical in the adult language, could also take the form *Go to get it*, and so an example

like (9a) cannot be counted as having either correct or incorrect tense marking in the subordinate clause, and are accordingly excluded from the scoring. Other examples, like (9d) *Adam put boot on looking for*, are too skewed—and the context too spare—to determine their structure. An example like (9e) *Laughing being cowboy*, though, may be unambiguously regarded as correct, since the corresponding subordinate clause in the adult language (*I'm laughing at being a cowboy*) also lacks tense. An utterance like *Laughing am cowboy* or *Laughing to be cowboy*, on the other hand, would be counted as incorrectly bearing tense. Such subordinate clauses are here called adjuncts, since they are not obligatorily selected by the matrix verb. The other very common kind of subordinate clause in Adam's early speech is the 'imperative complement', as in (9f) *Look birdie fly*. The adult language could realize the imperative complement as *Look at the birdie fly* or as *Look at the birdie flying*. However, because each of the adult variants lacks tense, an utterance like (9d) is counted as being correctly inflected. An utterance like *Look at the birdie flew* or *Look at the birdie to fly*, though, would be counted as incorrectly bearing tense. A sentence like (9g) *Gon (t)a ride dat* is regarded as containing a subordinate clause, since the semi-auxiliary *go* in the adult language obligatorily selects an infinitival *to* clause, and so if infinitival *to* were missing from such a clause (*Going ride that*), it would indicate that the T projection was missing from the subordinate clause. Such clauses are called 'to complements'. Another matrix verb that selects a *to* complement—which Adam uses most frequently—is *want*, as in (9b) *Want sit down*; here, infinitival *to* is missing from the embedded clause, and so the inflection of the embedded clause is counted as an error.

The general methodology, then, is to count all and only those subordinate clauses whose tense properties—i.e., whether or not T bears [\pm past] features or is headed by infinitival *to*—are unambiguously determined by the matrix predicate. Matrix predicates that may ambiguously select more than one complement type (with respect to tense) are counted as 'other complements' in the tables, but are not scored.

Results

It is particularly interesting that, from the moment that subordinate clauses begin to appear in Adam's speech, several different kinds appear at the same time. There seems not to be a period in which one kind of subordinate clause is learned, then another, and so on. An exception, however, is the class of unambiguously finite complements, which appear only rarely; for example, (9c) *Go belong* (which Adam's mother interprets as *Go where it belongs*).

Table 1 summarizes tense marking in Adam's subordinate clauses. Many of his first *to* complements are correctly inflected for tense:

- (10) a. Gon (t)a take suitcase out (Adam 2;4.3)
 b. Gon (t)a ride dat (Adam 2;6.3)
 c. Want to ride truck (Adam 2;6.3)
 d. Want to ride dat (Adam 2;6.3)
 e. Want to sit down (Adam 2;6.3)
 f. No wan(t) (t)a sit dere (Adam 2;6.3)
 g. Mommy want to ride dat (Adam 2;6.3)

These data initially suggest that the Truncation Hypothesis best characterizes the presence of tense in subordinate clauses. The main clauses of (10a,b,f,g), in particular, lack tense, while their subordinate clauses do have tense. However, after 2;6.3, Adam's performance on embedded tense declines significantly, so that by 2;11.0, he correctly inflects only one of his 55 *to* complements. Some examples are given below:

- (11) a. Do you want me get in (Adam 2;11.0)
 b. Do want he walk (Adam 2;11.0)
 c. Do you like come in with me (Adam 2;11.0)
 d. I going play baseball (Adam 2;11.0)
 e. I going swallow it (Adam 2;11.0)
 f. She going buy another one (Adam 2;11.0)
 g. I going drink it all up (Adam 2;11.0)
 h. Going take a wheels off (Adam 2;11.0)
 i. I going turn hot water on (Adam 2;11.0)

With respect to matrix clauses, Adam is still firmly in the OI stage, having dozens of such tokens, some of which are listed below:

- (12) a. Wheel turn? (Adam 2;11.0)
 b. Cowboy hat shoot with me (Adam 2;11.0)
 c. I say soxy soccer (Adam 2;11.0)
 d. Wait for grapefruit (Adam 2;11.0)
 e. Come with you (Adam 2;11.0)
 f. She need some (Adam 2;11.0)
 g. Make you very best (Adam 2;11.0)
 h. I spill it again (Adam 2;11.0)

This situation is very familiar from first language acquisition: the child initially appears to know the adult forms, but later begins making mistakes, appearing to regress. An explanation for Adam's initial, apparent knowledge of how to inflect *to* complements is suggested by the fact that the majority of the matrix verbs are *want* (with the occasional *go*). It is very likely that Adam, who was

only beginning to use embedded clauses, did not know that *to* is the infinitival marker, and instead analyzed *want to* and *going to* as single lexical items—not an unlikely possibility, given that these words are frequently contracted to *wanna* and *gonna* in adult English, and also even in Adam's speech in (10a,b,f) above. Further evidence for this misanalysis is that Adam acquired imperative complements at the same time, exemplified below:

- (13) a. Look birdie fly (Adam 2;5.12)
 b. Let Adam write (Adam 2;6.17)
 c. Let Adam cut it (Adam 2;6.17)
 d. Let Adam do it (Adam 2;6.17)
 e. Stop playing? (Adam 2;6.17)
 f. Let me ride dat (Adam 2;6.17)
 g. Let Adam finish (Adam 2;6.17)
 h. Oh let Adam tear it (Adam 2;6.17)
 i. Let Adam play busy bulldozer (Adam 2;6.17)

As imperative verbs select bare infinitival complements (rather than *to* infinitival complements), it is possible that, at around 2;6—when Adam appeared to inflect all embedded clauses correctly—he was applying a rule that required all embedded clauses to contain bare infinitivals, with the unanalyzed matrix verbs *wanna* and *gonna* falling perfectly into this schema. Once these verbs were analyzed into verb-plus-infinitival-*to* constructions (around 2;7.14), Adam's performance on *to* complements declined considerably. As shown above in (11), most of Adam's errors are caused by his producing bare verb complements for *want* and *go*, as he likewise produces for complements of imperative *let* in (13) (though correctly for that class of verbs). At this stage, then, Adam clearly has trouble distinguishing the several classes of embedded clauses.

These facts provide dramatic support for the Optional Tense Hypothesis. As was explained in the previous section, the Truncation Hypothesis does not predict that there will be tense errors in embedded clauses, since the non-appearance of tense in OIs is analyzed as clausal truncation. The Optional Tense Hypothesis, though, correctly predicts that embedded clauses will show the same proportion of missing T projections in the OI stage as do main clauses.

Although the cooccurrence (or not) of overt complementizers (e.g., *that*) with tense would yield further significant insight into the correctness of either the Truncation Hypothesis or the Optional Tense Hypothesis, overt complementizers unfortunately do not appear in any of the early data examined here. However, sentences like (11d-i) provide additional support for the Optional Tense Hypothesis, since these sentences have tense missing from both their main and embedded clauses—a distribution that is not predicted by the

Truncation Hypothesis, in which main clauses may lack tense while embedded clauses are expected to be intact. In contrast, the Optional Tense Hypothesis makes the broader (but correct) prediction that the tense projection may be absent generally.

Table 2 summarizes tense marking in Sarah's subordinate clauses. Although Sarah does not produce as many embedded clauses as does Adam, the data from Sarah corroborate the conclusions drawn above. Sarah most frequently produces *to* complements, like the following:

- (14)
- | | | |
|----|-------------------------------|----------------|
| a. | I wanna waid [= <i>ride</i>] | (Sarah 2;3.7) |
| b. | I wanna ride my horse | (Sarah 2;3.7) |
| c. | I want do that | (Sarah 2;7.28) |
| d. | Want do again | (Sarah 2;7.28) |
| e. | Want go read it | (Sarah 2;7.28) |

Although she has only two of them, Sarah's first *to* complements, in (14a,b), appear to be correctly inflected. However, like Adam's speech in (10a,b,f) above, the infinitival *to* has been contracted with the verb, suggesting that Sarah has likewise regarded *wanna* as a single lexical item, rather than as a contraction of *want to*. Like Adam, her matrix verb selecting a *to* complement is nearly always *want*, with an occasional exception like *come* (15f):

- (15)
- | | | |
|----|--------------------------|----------------|
| a. | I want have some in dere | (Sarah 2;8.25) |
| b. | I want put my | (Sarah 2;8.25) |
| c. | I want fit in there | (Sarah 2;8.25) |
| d. | I want see something | (Sarah 2;8.25) |
| e. | Want see | (Sarah 2;8.25) |
| f. | I come get you | (Sarah 2;8.25) |
| g. | I want see Romper Room | (Sarah 2;9.29) |
| h. | Want see Measles | (Sarah 2;9.29) |
| i. | I want do it | (Sarah 2;9.29) |
| j. | I want hear it | (Sarah 2;9.29) |
| k. | I want play record | (Sarah 2;9.29) |
| l. | I want buy two them | (Sarah 2;9.29) |

Note, importantly, that Sarah—like Adam—is still in the matrix OI stage when she begins producing embedded clauses:

- | | | | |
|------|----|-------------------|-----------------|
| (16) | a. | Her get my book | (Sarah 2;9.6) |
| | b. | I write crayon | (Sarah 2;9.6) |
| | c. | Just walk? | (Sarah 2;9.6) |
| | d. | He park here | (Sarah 2;9.6) |
| | e. | I pinch you | (Sarah 2;9.6) |
| | f. | I hurt again | (Sarah 2;9.6) |
| | g. | Her pour like dis | (Sarah 2;9.6) |
| | h. | Taste good | (Sarah 2;9.6) |
| | i. | Go square | (Sarah 2;9.29) |
| | j. | Her make pancakes | (Sarah 2;9.29) |
| | k. | Say moo | (Sarah 2;9.29) |
| | l. | He tickle a feet | (Sarah 2;9.29) |
| | m. | Sit potty | (Sarah 2;9.29) |
| | n. | I fall off | (Sarah 2;10.11) |
| | o. | I play | (Sarah 2;10.11) |
| | p. | I break it | (Sarah 2;10.11) |

The fact that Sarah's OI stage characterizes not only main clauses, but also her embedded clauses, argues in favor of the Optional Tense Hypothesis, and against a strong version of the Truncation Hypothesis.

Discussion

The Optional Tense Hypothesis predicts that tense may be freely absent from both main and embedded clauses, whereas the Truncation Hypothesis predicts that the tense projection is merely truncated along with other, higher functional projections. Given that proposition-taking verbs are generally assumed to subcategorize for full clausal complements like CP and IP, a strong version of the Truncation Hypothesis would predict that embedded clauses should retain their tense projection. The speech of Adam and Sarah shows that these children in fact omit tense across the board, and do not merely truncate the initial projections of their main clauses. This observation argues strongly for the Optional Tense Hypothesis.

A weaker version of the Truncation Hypothesis could be maintained if it were supposed that not merely root clauses, but also embedded clauses, could have their topmost projections truncated. However, this would entail loosening the isomorphism between semantics and syntax that is achieved by compelling proposition-taking verbs to select the maximal projection CP (or IP, for exceptional case-marking verbs). Such an analysis would fail to capture the intuition that the primary task facing the child in language acquisition is not the semantics of language, but rather the considerably more exception-ridden form of language, i.e., its syntax. Exchanges like ones in (17) show that children's

comprehension of embedded clauses is perfect, long before they can produce such clauses with correct tense marking themselves:

- (17) a. **Investigator:** Do you want to play with them?
Child: Yeah. (Adam 2;3.4)
- b. **Mother:** Wouldn't you like to pick these up?
Child: No. (Adam 2;3.4)
- c. **Mother:** What do you want me to do with the book?
Child: Read. (Eve 1;6)
- d. **Mother:** You gonna play music for us?
Child: Yep. (Eve 1;6)
- e. **Mother:** Santa Claus going to bring you another Bobo?
Child: [*nods*] (Sarah 2;3.22)
- f. **Mother:** Want me to take it away from you?
Child: No. (Sarah 2;3.26)

Facts like these are not surprising, particularly when one considers that general cognitive ability is acquired well before language ability, and extremely complex lexical properties are often acquired upon initial exposure (Chomsky 1995: 15). Empirically and theoretically, then, the Optional Tense Hypothesis offers the best account of embedded optional infinitives.

1. *Tense in Adam's subordinate clauses*

Adjuncts			Imperative complements			to complements			Finite complements			Other complements
Correct	Total	% Correct	Correct	Total	% Correct	Correct	Total	% Correct	Correct	Total	% Correct	(num)
0	0	-	0	0	-	1	6	17	0	1	0	4
0	0	-	1	1	100	0	2	0	0	1	0	6
0	0	-	1	1	100	1	1	100	0	2	0	2
0	0	-	0	0	-	0	0	-	0	0	-	0
0	0	-	0	0	-	0	0	-	0	0	-	0
1	1	100	1	1	100	0	0	-	0	0	-	2
1	1	100	0	0	-	6	6	100	0	0	-	1
3	3	100	8	8	100	26	31	84	0	0	-	2
0	2	0	8	8	100	9	10	90	0	0	-	4
2	2	100	19	19	100	0	5	0	0	0	-	3
0	0	-	26	26	100	1	9	11	0	1	0	5
0	0	-	29	32	91	0	5	0	0	0	-	0
0	0	-	6	6	100	1	5	20	0	0	-	2
0	0	-	4	4	100	2	8	25	0	0	-	0
0	0	-	7	7	100	1	16	6	0	0	-	3
0	0	-	16	16	100	4	26	15	0	1	0	0
1	1	100	25	25	100	1	55	2	0	1	0	2

Note

* Many thanks to Ken Wexler and the participants of WECOL 1996 for their comments on this paper. This research was supported by the Social Sciences and Humanities Research Council of Canada (grant 752-96-0602).

Bibliography

- Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (eds), *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press.
- (1995). Language and nature. *Mind: A Quarterly Review of Philosophy* 104(413), 1-61.
- MacWhinney, B. (1995). *The CHILDES Project: Tools for Analyzing Talk*. Second edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- & Snow, C. (1992). Tools for analyzing child language corpora. In A. W. Mackie, T. K. McAuley & C. Simmons (eds), *For Henry Kucera: Studies in Slavic Philology and Computational Linguistics*. Ann Arbor: Michigan Slavic Publications.
- (1990). The Child Language Data Exchange System: An Update. *Journal of Child Language* 17(2), 457-472.
- (1985). The Child Language Data Exchange System. *Journal of Child Language* 12(2), 271-295.
- Poeppl, D. & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language* 69(1), 1-33.
- Rizzi, L. (1994a). Early null subjects and root null subjects. In T. Hoekstra & B. D. Schwartz (eds), *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the 1991 GLOW Workshops*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- (1994b). Some notes on linguistic theory and language development: the case of root infinitives. *Language Acquisition: A Journal of Developmental Linguistics* 3(4), 371-393.
- Schütze, C. T., Ganger, J. B. & Broihier, K. (eds) (1995). *Papers on language processing and acquisition*. (MIT Working Papers in Linguistics, vol. 26.) Cambridge, MA: MIT Working Papers in Linguistics.
- Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot & N. Hornstein (eds.), *Verb movement*. Cambridge: C.U.P.

On "Fewest Steps"

E.G. Ruys

Utrecht University

This paper addresses some empirical and theoretical properties of economy constraints as defined in Chomsky 1995, chapter 4. It argues that three such constraints (Procrastinate, Fewest Steps, "No Redundant Features") have questionable theoretical properties. All are violable, without their interaction being properly defined; the latter two are "global" in such a way as to be incompatible with the overall framework; Procrastinate, furthermore, is underdetermined by the empirical data; and "NRF" appears to defy consistent definition altogether. It is proposed to redefine the Fewest Steps constraint in such a way that the effects of the other two can be derived from this constraint, without the accompanying theoretical problems. This move will also be seen to resolve some serious empirical problems associated with these constraints, which lie in the areas of expletive-placement in English and Subject-Object asymmetries in overt Accusative-checking languages.

1. Introduction

Recent developments in the Minimalist framework have shown a move away from violable, global economy constraints on derivations, in favor of inviolable, local constraints on movement (see Chomsky 1995, Chapter 4; hereafter: Chapter 4). This leads to a considerable reduction in computational complexity within the grammar, as well as providing a principled solution to complex questions arising from the interaction of multiple violable constraints.

Nevertheless, some violable constraints have as yet appeared irreducible, notably Procrastinate, Fewest Steps and Chomsky's "No Redundant Features" constraint; the latter two remaining not only violable, but global as well. In this paper, I will argue that the effects of both Procrastinate and "No Redundant Features" can be derived from Fewest Steps, if properly defined. This leaves us with just one violable, global constraint: a highly desirable result from a conceptual standpoint. Also, I will argue that this reduction has considerable empirical advantages.

I will start with a typology of (economy) constraints (section 2). I will then discuss some empirical and conceptual problems in Chapter 4, and propose a solution. These problems concern Chomsky's characterization of Procrastinate, his analysis of *there* and *it* insertion, and his "No Redundant Features" condition; the solution comes from a refinement of Fewest Steps.

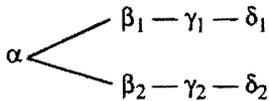
2. Constraint Types

For the purposes of our discussion, it is necessary to distinguish the following types of constraints: relative constraints vs. absolute constraints, and local vs. global constraints.

I will call a constraint *relative* (or “violable”) when it excludes a derivation (or a step in a derivation) iff the reference set of that derivation (or step) contains an alternative which violates the constraint to a lesser degree. A typical example is Last Resort. Suppose that for any given derivation, the reference set of that derivation with respect to Last Resort is the set of converging derivations with the same underlying numeration. Then Last Resort rules out a derivation A iff A requires a certain operation α , and there exists a convergent derivation B with the same numeration as A which does not require α . If there is no such alternative, however, then A and its operation α are not excluded by Last Resort.

An *absolute* constraint, on the other hand, excludes certain derivations *irrespective as to whether an alternative derivation is available*. Since a derivation that violates an absolute constraint is ruled out without being compared with possible alternative derivations, it is not necessary to provide a definition of Reference Sets for such constraints.

As for the distinction between “local” and “global” constraints: suppose there is a derivation (whose length may be zero) up to point α , and from α there is a choice between the continuations $\beta_1 - \gamma_1 - \delta_1$ and $\beta_2 - \gamma_2 - \delta_2$:



I will call a constraint “local” if it weighs the costs of β_1 and β_2 and chooses on the basis of these costs, *irrespective of the costs of $\gamma_1 - \delta_1$ and $\gamma_2 - \delta_2$* . A “global” constraint, on the other hand, computes the sums (or, conceivably, some other function) of the costs of $\beta_1 - \gamma_1 - \delta_1$ and of $\beta_2 - \gamma_2 - \delta_2$ and chooses between β_1 and β_2 accordingly. When a constraint is local it is possible in principle that it favors a derivation which in the end, from a global perspective, comes out as more costly.¹

In Chapter 4, Chomsky redefines most existing economy constraints. While most constraints in previous frameworks (see e.g. Chomsky 1991, 1993) were relative (violable) and all were global, most constraints in Chapter 4 are absolute, and some are argued to be local. Chomsky provides an absolute implementation of most economy-constraints by incorporating them into the definition of the movement-transformation. In that way, other options simply cannot be derived and cannot be taken into consideration. In OT-type terminology: most constraints are reinterpreted as properties of GEN. Once a condition is part of the definition of move, it is also local, since move does not create derivations in one fell swoop but creates them “step by step.” Hence, most conditions become not only absolute, but local as well.

Chomsky notes several reasons for preferring absolute, local constraints over relative and global ones, most having to do with a reduction of computational

complexity. In order to judge a derivation w.r.t. an *absolute* constraint it is no longer necessary to compare a derivation (or a step) with a possibly infinite set of alternative derivations (or steps). Furthermore, in order to judge a derivation w.r.t. a *local* (relative) constraint, it is no longer necessary to compare a set of sequences of steps; one need only compare a set of single steps. Also, Reference Sets for local constraints will presumably be smaller than reference sets for global constraints, since one compares only those derivations that are identical up to the point of the operation being judged (i.e. those derivations that have α in common). Finally, Chomsky points out that the number of available alternatives will decrease as the derivation progresses.

In addition, there is one clearly identifiable problem associated with the use of violable constraints. Consider a situation where derivation α is costly w.r.t. condition A, but inexpensive w.r.t. condition B, whereas derivation β is inexpensive for A but expensive for B. Now in order to decide which derivation is cheaper, one may be forced to add up different, incommensurable "costs."² The complex questions that arise in such a situation can be resolved in two ways. One option is to formulate a general theory of constraint interaction, as in Optimality Theory (see e.g. McCarthy & Prince 1994). The other option, pursued here, is to allow only absolute economy constraints (or at most one relative constraint).

Finally, of course, there may be empirical arguments in favor of absolute and local constraints; I shall consider at least one such argument below.

The remainder of this paper is organized as follows. First, in section 3 I will discuss Chomsky's analysis of *there* and *it* insertion. The analysis of *there* insertion is based on a crucially local view of Procrastinate. I will discuss a serious empirical problem for this analysis. I will then show that this problem can be solved if we keep Chomsky's analysis virtually intact, but make use of a (partly local) version of Fewest Steps, instead of Procrastinate. This seems to be a step back, since we then need two global, relative constraints: Fewest Steps, as well as Procrastinate. But in section 4, I will show that we can derive Procrastinate from Fewest Steps; this leaves us with just the one relative, partly global constraint. Finally, in section 5, I will discuss Chomsky's analysis of NP-to-Case movement. I will show, firstly, that this analysis gives independent evidence for Fewest Steps. Furthermore, I will discuss Chomsky's "No Redundant Features" constraint. I will show that this constraint is global and violable (hence should be avoided on general grounds) and, furthermore, does not work as stated, and finally, that its effects (at least those discussed here) can *also* be derived from Fewest Steps. So, rather than being faced with three global and violable constraints (partly ineffective), we are left with just one such constraint, Fewest Steps, which is effective for the data discussed here.

3. *There and It: Procrastinate or Fewest Steps*

Chomsky discusses the following pair:

- (1) a. There seems t to be someone in the room.
- b.* There seems someone to be t in the room.

Why is (1a) in and (1b) out? Chomsky's analysis runs as follows. Both derivations have (2) in common as a subderivation.

- (2) [_{I'} [_{INFL}[EPP] to] [_{VP} be someone in the room]]

In (2), EPP must be checked on INFL. There are two possible ways of achieving this: move *someone* to Spec,IP, or insert *there*. If we move *someone*, we derive (1b) through the steps indicated in (3):

- (3) 1. [_{IP} someone [_{INFL} to] [_{VP} be _t_{someone} in the room]]
 2. [_{IP} there INFL seems [_{IP} someone [_{INFL} to] [_{VP} be _t_{someone} in the room]]
 3. [_{IP} there FF(someone) INFL seems [_{IP} someone _t_{FF(someone)} [_{INFL} to] [_{VP} be _t_{someone} in the room]]

After movement of *someone* (step 1) and further derivation, *there* must be inserted to check EPP on matrix-INFL and deplete the numeration (step 2). Finally, after Spell-Out the formal features associated with *someone*, FF(*someone*), move to matrix INFL, checking NOM and ϕ -features (step 3). This derivation results in the ill-formed (1b). The derivation has one overt movement (violation of Procrastinate) and one covert movement. Now consider the derivation of the well-formed (1a). Again, we start from (2), but we proceed as in (4):

- (4) 1. [_{IP} there [_{INFL} to] [_{VP} be someone in the room]]
 2. [_{IP} there INFL seems [_{IP} _t_{there} [_{INFL} to] [_{VP} be someone in the room]]
 3. [_{IP} there FF(someone) INFL seems [_{IP} _t_{there} [_{INFL} to] [_{VP} be someone _t_{FF(someone)} in the room]]

Step 1 inserts *there*. Subsequently, matrix EPP must be checked and the Minimal Link Constraint (MLC) requires that this be done through movement of *there* (step 2). Finally, after Spell-Out, FF(*someone*) move to matrix-INFL, checking NOM and ϕ (step 3). This derivation has one overt movement (violation of Procrastinate) and one covert movement, as well. The result is well-formed (1a).

Given that both derivations have the same number of movements and violations of Procrastinate, why should (1b) be out and (1a) in? Chomsky's answer is that (1b) violates Procrastinate earlier in the derivation. At the common point in the derivation, (2), there is a choice between moving *someone*, and inserting *there*. Now assume that Procrastinate is a local constraint. It then prefers insertion over movement at this choice point, and the fact that insertion will inevitably be followed by a Procrastinate violation further down the line is "invisible." Thus, although both derivations are equally costly from a global perspective, (1a) is preferred because the relevant constraint (Procrastinate) operates on a purely local basis. (We have then an empirical argument for a local constraint.)

Now note, first of all, that we can replace “Procrastinate” with “Fewest Steps,” in the analysis presented above, and have the same result fall out in the same way. At the choice point (2), we have a choice between insertion and movement. Given the fact that insertion is “costless,” whereas movement bears a cost, Fewest Steps will prefer insertion, and we derive (1a). (1b) is equally costly w.r.t. Fewest Steps from a global perspective (both have two movement operations, one overt and one covert), but by assuming that Fewest Steps is a purely local constraint, we derive that (1a) is preferred over (1b).

Although “Local Fewest Steps” and “Local Procrastinate” give the same result in this case, I will argue next that “Fewest Steps” is preferable when *it*-expletives are taken into account. Consider first the examples in (5).

- (5) a. it_i seems [IP t_i to appear to John [CP that ...]]
 b. *it seems [IP John $_j$ to appear to t_j [CP that ...]]

The contrast in (5) shows that *it*-insertion behaves just like *there*-insertion in (1). Both derivations converge, but (5a) is locally cheaper because at the choice point where EPP must be checked on the embedded INFL, a choice has been made for merger instead of movement. Again, both derivations are equally expensive globally; only local Procrastinate or Fewest Steps can make the correct distinction. But the situation is reversed in (6):

- (6) a. *it seems [CP that t was told John [CP that ...]]
 b. it seems [CP that John was told t [CP that ...]]

The contrast between (6a) and (6b) is problematic. Their common substructure is:

- (7) [I' INFL[EPP] was told John [CP that ...]]

At this point, there is a choice: move *John*, or insert *it*. If we move *John* (violating local Procrastinate or Fewest Steps), the derivation proceeds as follows:

- (8) 1. [IP John INFL was told t_{John} [CP that ...]]
 2. [I' INFL[EPP] seems [CP that [IP John INFL was told t_{John} [CP that ...]]]]
 3. [IP it INFL seems [CP that [IP John INFL was told t_{John} [CP that ...]]]]

Followed by Spell-out. This derives the well-formed (6b). Alternatively, we can start from (7) and insert *it* (obeying local Procrastinate or Fewest Steps):

- (9) 1. [IP it INFL was told John [CP that ...]]
 2. [I' INFL[EPP] seems [CP that it INFL was told John [CP that ...]]]]
 3. [IP it INFL seems [IP that t_{it} INFL was told John [CP that ...]]]]

This is followed by Spell-Out. After Spell-Out, we can move FF(*John*) to matrix-INFL, checking NOM on both. The result is:

- (10) [_{IP} it FF(*John*) INFL seems [_{IP} that t_{it} INFL was told John t_{FF(*John*)} [_{CP} that ...]]]

This derivation converges, and derives the ill-formed (6a). On the basis of the preceding discussion of *there*-insertion, then, we should expect (6a) to be favored over (6b), since the derivation of (6a) is locally more economical.

Chapter 4 attempts to solve this problem by claiming that (6a) does not converge: the Case features on matrix INFL as well as the Case features on *John* remain unchecked — in that case, (6a) is not contained in the Reference Set for (6b) and (6b) is not blocked. But we have seen that this solution is not correct. Note, that in the derivation of the well-formed (1a), Case features on matrix-INFL and on *someone* also remain unchecked in overt syntax. Nevertheless, the derivation converges, because FF(*someone*) can move to matrix-INFL at LF, so that both features are checked. Nothing prevents the same operation from applying in (6a) (see (10)).

So, in sum, local Procrastinate (or Fewest Steps) prefers derivation (7)-(9)-(10) over derivation (7)-(8), hence favors ill-formed (6b) over well-formed (6a). I shall skip discussion of various implausible ways out, and move straight to the solution I want to propose. If we compare ill-formed (7)-(9)-(10) to well-formed (7)-(8), it turns out that the ill-formed derivation has more movement operations: it requires a covert movement of FF(*John*), whereas (7)-(8) requires no covert movement. Thus, although (7)-(9)-(10) is less expensive from a local perspective, it is more expensive from the perspective of a *global* Fewest Steps constraint. I propose, then, a Fewest Steps constraint which is partly local, and partly global:

(11) **Fewest Steps (FS)**

Given two derivations, choose the one with the fewest steps; in the case of a tie, choose the one which is locally cheaper (which takes the cheaper option at the choice point where the derivations diverge).

This gives the same result for *there*-insertion in (1) and *it*-insertion in (5): both derivations are globally equally expensive, hence local economy decides. But in the case of *it*-insertion in (6), global Fewest Steps prefers (6b), as desired.

One may argue that Fewest Steps as we have now defined it is actually two constraints, not one. The point is hardly worth discussing, apart from the fact that, if these are two conditions, their interrelation and interaction is well-defined, and does not cause the type of problems associated with the interaction of violable constraints discussed in section 2.

This reanalysis allows us to choose between Fewest Steps and Procrastinate as the constraint to be held responsible for the data discussed in this section. Whereas Fewest Steps allows for a natural (global) extension in order to capture the data in

(6), a similar extension of Procrastinate will not work: (6a) and (6b) each violate Procrastinate once, hence if we were to add a global Procrastinate clause, it would be inoperative and the local clause would incorrectly choose (6a).

This section leaves us with an addition to our rule-set: a global clause in the Fewest Steps constraint. In the next section however, I will argue that adopting local Fewest Steps allows us to eliminate Procrastinate altogether.

4. *Deriving Procrastinate*

The Procrastinate principle, introduced in Chomsky 1991, is a relative (violable) constraint. A derivation that has an overt operation is excluded by Procrastinate iff the Reference Set of that derivation contains a (converging) alternative derivation that has fewer overt operations. As pointed out in section 2, we want to avoid such constraints. In particular, since we have other relative constraints (such as Fewest Steps) next to Procrastinate, we face the intricate problem of controlling the interaction of several relative constraints. We can avoid this problem by eliminating Procrastinate.

There is a second reason to want to derive Procrastinate from other constraints. Observe that Procrastinate is the exact mirror-image of Pesetsky's (1989) Earliness principle, which states that movement operations must take place as early as possible. For instance, whereas Procrastinate might explain the difference between Wh-movement in English and Chinese by stating that Wh-movement *can* take place post-Spell-Out in Chinese, hence *must* take place post-Spell-Out, but cannot take place post-Spell-Out in English, hence is bound to occur early, Earliness says that Wh-movement *can*, hence *must* take place early in English, but cannot be early, hence is condemned to occur late in Chinese. We could build a mirror-image of the Chapter 4 model in which Earliness is operative instead of Procrastinate, and in which weak features have the property that they cannot be checked pre-Spell-Out. This implies that Procrastinate as a principle of grammar is underdetermined by the facts. It would be interesting, therefore, if it could be shown that the principle that distinguishes overt from covert movement (Earliness or Procrastinate) *must* be Procrastinate. We achieve this aim if we can derive Procrastinate as an effect of other principles.

In Chapter 4 Chomsky suggests that perhaps, Procrastinate need not be stipulated as an (axiomatic) condition, but may be derivable from other constraints. The idea is roughly this: when a feature moves early, i.e. pre-Spell-Out, the PF-component demands that Pied-Piping takes place: the PF-component cannot interpret "unbound features." No Pied-Piping is necessary after Spell-Out has occurred. As a consequence, "more material" is moved during overt movement than during covert movement, so that it is natural to suppose that overt movement is more expensive than covert movement; economy then prefers covert movement if possible.

This account is incomplete unless we can determine exactly which economy constraint is involved here. There are two possible views of the matter, depending on the exact properties of the Pied Piping operation.

The first option builds on Chomsky's assumption that overt movement creates (at least) two chains: the chain(s) of the formal features, one of which has been attracted, and the chain of material being Pied-Piped in the process. This option has two disadvantages. First, it implies a complication in the definition of the movement transformation, since it must now be able to create different numbers of chains in a single operation (creating the chain of the Pied Piped material cannot be a separate operation, since this material is not itself attracted). Secondly, in order to derive Procrastinate from this, we must adopt a global, relative economy constraint which counts the number of chains in a derivation. The constraint must be global, as it compares the number of chains resulting from two possible operations that apply at different points in the derivation. This interpretation of Procrastinate takes us further from our goal, in that it provides independent evidence for a global, relative constraint of the type which Chapter 4 seeks to avoid.

A second option is to assume that Attract-F always creates one chain, which optionally consists either of the formal features, or of both the formal features and the pied-piped material. This option allows attract-F to remain relatively simple. But in order to derive Procrastinate from economy we must now assume the existence of an additional global, relative economy constraint on top of those discussed earlier, namely one which measures the "weight" or "amount" of material that is moved by one application of move-F, and compares it to the "amount moved" by a later application of move-F. Chomsky formulates such a condition: "[the attracted feature] F carries along just enough material for convergence." Independent evidence for this condition should come from restrictions on (traditional) Pied Piping, but it rather creates problems in that area, as Pied Piping allows of a fair amount of optionality.

Note, furthermore, that both these options derive Procrastinate from a *global* constraint, so that they are inconsistent with Chomsky's *local* view of Procrastinate as it supposedly applies to *there* insertion (see previous section).

We can avoid all talk of "numbers of chains created by a single application of attract-F," or of "amounts of material moved by Attract-F," and derive Procrastinate from a purely local constraint, along the following lines.

Assume first that Fewest Steps economy (11) needs to consider only those operations that are *potentially* superfluous; i.e., assume the following dichotomy:

Grammatical operations	
counted by FS	not counted by FS
Attract	Select
	Merge
	Spell-Out

The intuition underlying this dichotomy is that selecting a certain lexical item from the numeration, merging structure K with some distinct K', and applying Spell-Out to a derivation must each apply once, and cannot apply more than once. An item in

the numeration must be selected once, and cannot be selected more than once, etc. Only the movement operation *Attract* can apply to a given item more than once, and potentially more often than necessary; hence FS considers only *Attract*.

If we assume that FS applies only to *Attract*, we immediately derive the correct relative order of operations in a derivation:

Procrastination: $\text{Select/Merge} \mid \text{Attract}_{\text{Strong F}} < \text{Spell-out} < \text{Attract}_{\text{Weak F}}$

To illustrate, assume a numeration with a certain number of lexical items with strong and weak features, and consider which operation must be applied at each given point in the derivation. Initially, there is no choice: *Select* (from the numeration) and *Merger* are the only options. Only after *Select* and *Merger* have applied a number of times, and have created a structure where a head H attracts a feature in its complement, do we have a choice: *Merger*, or *Attract-F*.

Because *Merger* is costless, whereas *Movement* is costly, local *Fewest Steps* (as defined in section 3) will always choose *Merger* at each specific choice point (even if *Movement* will have to occur sooner or later anyway). Hence, a choice for *Movement* will be made only when this is necessary for convergence: when H has a strong feature, and merger of HP with a head G will lead to a proscribed structure in which a strong feature is contained in an embedded projection. So, movement for weak feature checking will always be postponed until there simply are no other options left; i.e., until the numeration has been depleted, and *Spell-Out* has occurred. Any derivation therefore must proceed in the following order: first strong features are checked and lexical items are merged, then *Spell-Out* is applied, and after that, weak features are checked. In this way, the relative ordering of weak feature checking after strong feature checking and depletion of the numeration falls out as an effect of local economy.³

In this section we have resolved the tension between two violable constraints, *Fewest Steps* and *Procrastinate*, by deriving one from the other. We have also strengthened the case for local economy, by deriving procrastination effects from the local clause of FS. The next section addresses a third relative constraint that plays a role in Chapter 4.

5. Case Checking, Fewest Steps and Redundant Features

Besides *Procrastinate* and *Fewest Steps*, Chapter 4 introduces one further global and violable constraint. A numeration underlying a syntactic derivation may not contain features which have no effect on output, i.e. on PF or LF: “ α enters the numeration only if it has an effect on output.” We will call this the *No Redundant Features Constraint* (NRF). In the present section we will show that NRF is global and violable (hence to be avoided on general grounds, as indicated above), as well as insufficiently well-defined to be reliably utilized. It will turn out that at least some of the effects attributed to NRF can be deduced, again from *Fewest Steps*. We will illustrate these issues by means of a single set of examples, involving Case-checking in an overt Accusative-checking language.

A model of grammar which analyses Case-assignment as movement to a Case-checking position must somehow ensure that the correlation between θ -features, Case-features, and overt NP-positions remain intact (the right NP in the right position has the right Case). Thus, well-formed Icelandic (12a) must be allowed but (12b) must be ruled out (for discussion of Case in Icelandic see e.g. Sigurðsson 1989, Jonas & Bobaljik 1993).

- (12) a. *María las bókina.* (Icelandic)
 [_{Subject} Mary-NOM] read [_{Object} the-book-ACC] 'Mary read the book'
 b. **Bókina las María*
 [_{Subject} the-book-ACC] read [_{Object} Mary-NOM] 'The book read Mary'

We start then with an Icelandic numeration with strong DP-features on both T and \underline{v} , a transitive verb, and two DPs, one (*bókina*) marked with accusative, one (*María*) with nominative Case. Initially, we have two options: merge NP_{NOM} in subject position (external θ -role) and NP_{ACC} in object position (internal θ -role), or vice versa. We want the former choice, (12a), to successfully converge, and the latter choice, (12b), to be ruled out.⁴

Let us start with the desired initial derivation:

- (13) [_{VP} SU_{NOM} [_v VB [_{VP} t_v OB_{ACC}]]]
 1 [_{VP} OB [_v SU_{NOM} [_v VB [_{VP} t_v t_{OB}]]]]
 2 [_T T [_{VP} OB [_v SU_{NOM} [_v VB [_{VP} t_v t_{OB}]]]]]]
 3 [_{TP} SU [_T T [_{VP} OB [_v t_{SU} [_v VB [_{VP} t_v t_{OB}]]]]]]]]

Steps 1 and 2 are forced. Moving SU to outer spec,vP in 1 would have resulted in a Case-mismatch, so OB moves to check the strong features on \underline{v} . Subsequent structure-building derives step 2. Moving SU (step 3) in order to check strong DP features of Tense correctly derives (12a) (after V2).

Note, incidentally, that the last step could have moved OB to Spec,TP instead, incorrectly deriving a structure with the object in Spec,TP. But this is ruled out by the global clause of Fewest Steps, since it requires that FF(SU) move covertly to check NOM. Thus, this derivation provides independent evidence for global FS.⁵

Consider now the unwanted derivation of (12b) in (14):

- (14) [_{VP} SU_{ACC} [_v VB [_{VP} t_v OB_{NOM}]]]
 1 [_{VP} SU [_v t_{SU} [_v VB [_{VP} t_v OB_{NOM}]]]]
 2 [_T T [_{VP} SU [_v t_{SU} [_v VB [_{VP} t_v OB_{NOM}]]]]]]
 3 [_{TP} SU [_T T [_{VP} t_{SU} [_v t_{SU} [_v VB [_{VP} t_v OB_{NOM}]]]]]]]]
 4 [_{TP} SU [_T [_T FF(OB) [_T VB T]] [_{VP} t_{SU} [_v t_{VB} [_{VP} t_v OB+t_{FF}(OB)]]]]]]

In step 1 we must move SU to outer Spec,vP, checking ACC and strong DP-features on \bar{y} (moving OB instead would lead to a mismatch of Case-features, hence terminate the derivation). Subsequent structure-building derives step 2. Now, strong DP-features on Tense must be checked, and MLC only allows SU to move (step 3). Covertly, FF(OB) move to Tense and NOM is checked. This derivation converges; all uninterpretable features have been checked. The result is ill-formed (12b).

Chapter 4 attributes the ill-formedness of the unwanted derivation resulting in (12b) to the fact that, given that derivation, the numeration that it proceeds from is too expensive, as it contains a superfluous strong DP feature on \bar{y} . For without that feature we would have arrived at the same PF and LF (with SU moving from Spec,vP to Spec,T in one swoop, without an intermediary landing in outer Spec,vP as in step 1). Let us consider this analysis more closely.

First of all, we see that NRF is not a local condition. Whether a feature will have an effect on output cannot be judged during the composition of the numeration. For whether a feature will have an effect depends on other choices that are made during the derivation. The intended effect of NRF then is to preclude those choices that deprive a feature of the chance of having an effect on output. In the present example, the strong DP-feature on \bar{y} is supposedly superfluous, but it does play a role in the well-formed derivation (13) which proceeds from the same numeration (without it, OB remains in situ, to the right of VP-boundary adverbials). It is only when the accusative DP is merged as external argument that the feature becomes a superfluous component of the numeration in hindsight. So NRF is not a local constraint. Also, it is a relative constraint; a feature can only be judged to be superfluous in the presence of alternative converging derivations that can do without it.

But the conceptual problems are actually far more serious. We must wonder through what comparison of derivations NRF can produce the effect that Chomsky attributes to it. Consider, in the abstract, a condition C which states that operation O may be applied only if O has an effect on output. How can we ascertain the existence of condition C? We can only ascertain the existence of C by observing that C prevents O from applying in certain cases. And we can only observe whether O has been prevented from applying, in cases where application of O will produce a different output than non-application of O. But in those exact cases, in which application of O yields a different output than non-application of O, C will not prevent O from applying. Hence we can never observe a case in which C has prevented O from applying, for in those cases application or non-application of O is unobservable. C therefore makes no predictions which can be tested from observation: a grammar that contains C yields the exact same PF-LF pairs as a grammar that does not contain C.

Conditions of this type are meaningless, then, unless we adjust them along the following lines. Reformulate condition C as follows: "operation O may be applied only if O has a class A output effect." Now suppose that O can have class A and class B output effects. We can then ascertain the existence of condition C by

observing that a class B effect of O occurs only in combination with a class A effect (whereas a class A effect can occur without a class B effect).⁶

Example (12b), e.g. cannot be ruled out by NRF without such a class A-B distinction. NRF might rule out its derivation in (14), but only if output (12b) can be obtained from some sparser numeration — a self-defeating strategy.

Let us see what happens if the numeration for (12b) has no strong DP-feature on \underline{y} . Given this numeration (no EPP on \underline{y} , otherwise the same as for (13), (14)), one option is to merge NP_{ACC} as OB and NP_{NOM} as SU. But this will never lead to (12b), as the distribution of θ -roles over NP_{ACC} and NP_{NOM} will then be different.

Another option is to start as in (14); the derivation then proceeds as follows:

- (15) 1 [T T [_{VP} SU_{ACC} [_V VB [_{VP} t_V OB_{NOM}]]]]
 2 [TP SU_{ACC} [T T [_{VP} t_{SU} [_V VB [_{VP} t_V OB_{NOM}]]]]

Step 2 is forced by MLC; the derivation terminates due to mismatch of Case features (ACC on SU against NOM on T). The exact outcome now depends on what is meant by "termination." Either there are no LF and PF at all (in which case we have not managed to derive the same output as in (14)); or (15) step 2 is the output, but with a "*" because of termination. In the latter case (15) differs from (14) in two respects: (15) has a "*", and (15) has unchecked Case-features which (14) does not have, as well as having fewer chains. If the first difference were the only difference, this would make for an interesting and acceptable version of a ("class B") "insufficient difference." We could then understand NRF as follows: a converging derivation is ruled out, if from the same numeration minus one feature we can derive an otherwise identical PF-LF pair, that differs at most in that it has a "*". But, as noted, there are other differences between (14) and (15) which would also have to be considered insufficient, and it is not clear how this can be justified. In particular, (15) has an uninterpretable unchecked NOM-feature on OB: we would have to consider this distinction between (14) and (15) insufficient. That is, we must assume that the EPP-feature on \underline{y} can be "redundant" even if it is necessary for convergence.

We can complicate matters further. We have assumed so far, that NRF rules out a derivation from numeration N_1 with feature F, if (virtually) the same PF-LF pair can be obtained from numeration N_2 which differs from N_1 *only* in that it lacks F. But we might also take into consideration other numerations that differ more radically from N_1 : this might lead to a different theory of class A / class B properties of output. In our example, an obvious possibility to consider is that we could compare derivation (14) and its numeration N_1 (with EPP on \underline{y}) with the derivation from a numeration N_2 without EPP on \underline{y} , in which the NP that bears ACC in N_1 , bears NOM Case, and vice versa. We can then derive an alternative for (14b), i.e. for (12b), in which the same NPs bear the same θ -roles, but in which

the subject bears NOM and the object bears ACC. Since \bar{y} in N_2 lacks EPP, this would lead to a sentence without overt object movement, hence indeed to (12b) as far as word order is concerned. The PF, however, would be different, because morphological overt Case marking would be distributed differently over the two NPs. After all, the incorrect distribution of overt Case marking is the reason we want to exclude (12b) in the first place. Now, if we ignore this distinction (as a class B effect) then we indeed manage to exclude (12b). Paradoxically, we then manage to explain the Icelandic intuition that Case marking is wrong in (12b) by stipulating that Case marking is the very effect that is ignored when we judge the derivation for (12b) in (14) under NRF. Also, we then call upon a much wider, relative comparison of derivations, since we can no longer use numerations (or simple functions of them) to determine Reference Sets for NRF.

In any case, it should be clear that NRF cannot be said to be well-defined without there being some theory of "class A" (sufficient) and "class B" (insufficient) distinctions among PF-LF pairs, and that the properties of such a theory are not immediately evident. Global Fewest Steps, on the other hand (which by now has considerable independent motivation) straightforwardly explains the distinction in (12): the desired derivation in (13) requires two cases of DP-movement, whereas the undesired derivation resulting in (14) requires three (one covert movement more). Whether all distinctions presumed to follow from NRF can be attributed to Fewest Steps remains to be established, however.⁷

In conclusion. While Chapter 4 abolishes most relative constraints and replaces them with absolute properties of Attract, at least three violable constraints remain: Procrastinate, "No Redundant Features," and Fewest Steps. Procrastinate is held responsible for the distinction between overt and covert movement, and for the *there*-insertion data discussed in section 3. I have proposed that Procrastinate can be done away with, if we assume that Chomsky's global Fewest Steps is also partly local. Such a constraint can account for *there*-insertion just as well, and for *it*-insertion even better (section 3); it can also be held responsible for the overt/covert distinction (section 4). There is empirical gain here, as well as conceptual gain on three points. Two constraints have been replaced by one. Our analysis of procrastination effects strengthens the empirical basis for a local conception of economy, in that Procrastinate is derived from the local clause of Fewest Steps. And by deriving Procrastinate from other conditions we provide evidence for the assumption that the overt/covert distinction must be viewed in terms of procrastination, rather than earliness. In fact, procrastination is now explained on the basis of the fact that derivations are directional, and proceed "from numeration to LF." Finally, in the last section, we have shown that there is some promise that Fewest Steps can also replace NRF, a constraint with questionable properties in various respects. While we started out with three violable constraints, whose interaction was undefined, we are now left with just one violable, partly local constraint, and it appears that there may be no need for a theory of constraint interaction in syntax.

For discussion and comments I am grateful to Peter Ackema, Peter Coopmans, Johan Kerstens, Ellen-Petra Kester, Tanya Reinhart, Maaïke Schoorlemmer, and Fred Weerman.

¹ Our definition differs from Chomsky's, in that Chomsky also considers a constraint "local" if it compares global costs, but these costs can be predicted locally through some precalculation. Such a constraint can be global in our terms. Conversely, a local constraint on our definition can be "global" in the sense that computationally, lookahead is still necessary (e.g. to determine whether a possible step will lead to convergence). So, we have two disjoint notions of "local/global"; unlike Chomsky we shall not conflate the two and consider only one.

² An example of this type of problem is the interaction of Fewest Steps and Shortest Link as discussed in Chomsky 1993. If a given movement operation can take few or many intermediate landing sites, then a lower number of operations (sub-movements) implies an increase in the length of links, and vice versa. Chomsky solved this specific problem, not by adopting a general theory of constraint interaction, but by ensuring that the example situation cannot arise (a movement operation which involves various sub-movements counts as one operation Form Chain). See Kitahara 1995, note 26 for another illustration of this type of problem.

³ Deriving Procrastinate in this manner has been considered before (see Zwart 1993), but makes little sense unless Fewest Steps is recognized as a local constraint. Note, that the relative ordering of Spell-Out w.r.t. Select/Merge and Strong feature checking must follow from independent principles: we assume that Spell-Out cannot occur unless the numeration is depleted, and cannot apply to a structure that still contains some strong features. Note further, that we can now do without the global, violable constraint that "F carries along just enough material for convergence."

⁴ Note, that we are comparing derivations with different meanings. This is not a problem since we assume, with Chomsky 1995, that Reference Sets are determined solely on the basis of numerations. We ignore the semantic oddity of (12b), which can easily be fixed; what is relevant is that (12b) has an accusative subject.

⁵ Chomsky suggests that Fewest Steps works locally in preventing this derivation, but, as we have defined "local constraints," this is incorrect (cf. footnote 1). In step 3 movement must occur. Locally, moving OB is just as expensive as moving SU. Only a global Fewest Steps will consider it relevant that moving OB will eventually lead to an extra movement operation.

⁶ A complex variety of such a class distinction plays a role in the work of Fox that underlies Chomsky's treatment of the QUANT feature (see Fox 1993).

⁷ Chomsky uses NRF in one other context: to prevent merger of +Wh[+strong] COMP after Spell-Out in English. Here, it appears that NRF *can* apply non-vacuously, because [+strong] is superfluous, but also obligatory. There

is no alternative derivation without [+strong], so that pre-Spell-Out merger and overt Wh-movement are the only option. However, here the applicability of NRF is provable only metalinguistically, not within the system, as there is no derivation without [+strong] that the system can use for comparison. See note 3 for our alternative.

6. References

- Chomsky, N. 1991. Some Notes on Economy of Derivation and Representation. In *Principles and Parameters in Comparative Grammar*, ed. R. Freidin, 417-454. Cambridge, Mass: MIT Press.
- Chomsky, N. 1993. A minimalist program for linguistic theory. In *The View from Building 20*, ed. K. Hale & S.J. Keyser, 1-52. Cambridge, Mass: MIT Press.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass: MIT Press.
- Epstein, S.D. 1992. Derivational Constraints on A'-Chain Formation. *Linguistic Inquiry* 23:235-259.
- Fox, D. 1993. Quantifier Scopepe in VP ellipsis. Ms MIT.
- Jonas, D. and J. Bobaljik. 1993. Specs for Subjects: The Role of TP in Icelandic. In *MIT Working Papers in Linguistics 18: Papers on Case and Agreement 1*, 59-98. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Kitahara, H. 1995. Target alpha: Deducing Strict Cyclicity from Derivational Economy. *Linguistic Inquiry* 26:47-77.
- McCarthy, J. and A. Prince. 1994. *The Emergence of the Unmarked*. Ms. U. Mass/Rutgers.
- Pesetsky, D. 1989. Language Particular Processes and the Earliness Principle. Paper presented at *GLOW*, Utrecht, The Netherlands.
- Reinhart, T. 1993. Wh-in-situ in the framework of the Minimalist Program. Ms. UU.
- Sigurðsson, H.A. 1989. Verbal Syntax and Case in Icelandic. Diss. University of Lund.
- Zwart, C. J.-W. 1993. *Dutch Syntax: a Minimalist Approach*. Diss., University of Groningen.

E.G. Ruys
 Utrecht Institute of Linguistics OTS / Utrecht University
 Trans 10
 NL-3512 JK Utrecht
 The Netherlands
 ruys@let.ruu.nl

ENGLISH NOMINALS WITHOUT OVERT NOUN HEADS:
A LEXICAL SHARING ANALYSIS*

Michael T. Wescoat

Osaka University

1. INTRODUCTION

In this study I will employ a novel view of syntactic tree structure employing a notion I call *Lexical Sharing* to analyze the sorts of intuitively 'headless' nominal expressions illustrated in (1):

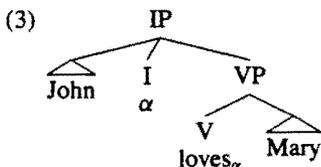
- (1) a. I've looked at *those*.
 b. I've spoken with {*many/most/few*}.
 c. A good time was had by {*all/both*}.
 d. I've thought of *two*.
 e. I commented on *John's*.

The intuition that the nominals in question are in some sense 'headless' can be simply articulated by the observation that each of the expressions highlighted in (1) could be followed by a common noun. What one can say with reasonable certainty about the structures in (1) is that the highlighted expressions are clearly DPs—the framework pioneered by Fukui (1986) and Abney (1987) will be adopted here. This is amply demonstrated by the fact that the highlighted phrases are all acceptable as objects of prepositions. Far less clear are issues concerning constituency and syntactic category affiliation of elements within the DP. Also troublesome is the issue of what sort of modifiers are acceptable in conjunction with the expressions in (1); for instance, (2) is acceptable only when the modifier's complement is present:¹

- (2) *Those ambitious *(enough to succeed)* will do well.

I shall show that the syntactic properties of forms like those in (1) are difficult to capture with phrase structure trees as they are traditionally conceived but fall out as natural consequences of the application of structures employing *Lexical Sharing*.

Elsewhere I introduce *Lexical Sharing* to treat such topics as English gerunds and *wh* constructions as well as Hindi noun incorporation (Wescoat 1994, 1995, 1996). The basic workings of this device are illustrated in (3):



Greek variables indicate sameness of lexical tokens; i.e., I and V in (3) are associated with the very same word, whence the term '*Lexical Sharing*'. This approach

obviously allows for individual words to have dual category status. Also, with the introduction of lexical sharing comes the need for a new view of left-to-right ordering among nodes; for instance, if I and V in (3a) are associated with the same word, *loves*, then there is no clear sense in which one of these nodes may precede the other. I reaxiomatize left-to-right ordering to allow for cases of this kind, and the resulting treatment of precedence will be crucial in sorting out a solution to the problem posed by (2).

This study will be structured in the following way. In section 2 I will examine the phenomenon illustrated in (2) in greater detail, showing how it resists analysis under standard assumptions about phrase structure. Section 3 then presents a reaxiomatization of phrase-structure trees that incorporates Lexical Sharing. This sets the stage for section 4, where I show how Lexical-Sharing makes the facts fall out without any ad hoc stipulations. Section 5 provides concluding remarks.

2. A PARADOX

Let us begin by examining in greater depth the unusual contrast in (2), which gives rise to a paradox under standard assumptions about the nature of phrase-structure trees. The observations offered in this section pertain to all the forms seen in (1), whether they feature a determiner, quantifier, or possessor; however, the form *those* from (1a) shall be used throughout this section as a representative for all the nominal types in (1). One finds that some attributive modifiers are acceptable, but not others, according to the pattern in (4) and (5):

- | | |
|--|-------------------------------------|
| (4) *those ambitious | <i>Simple APs</i> |
| (5) a. those ambitious enough to succeed | <i>APs with complements</i> |
| b. those {interested/willing} | <i>Past and present participles</i> |
| c. those {with/that have} ambition | <i>PPs and relative clauses</i> |

Predicting this contrast proves more challenging than might at first be expected.

Now, there is a rather obvious generalization to be made about the distinction between the modifiers in (4) and (5). It becomes visible if one considers word order in a normal DP. An attributive modifier may occur with a lone determiner if and only if it *may follow* a common noun, i.e., if and only if the corresponding (b) example in (6)–(9) is grammatical:

- (6) a. ambitious people
- b. *people ambitious
- (7) a. *ambitious enough to succeed people
- b. people ambitious enough to succeed
- (8) a. those {interested/willing} individuals
- b. those individuals {interested/willing}
- (9) a. *{with/that have} ambition people
- b. people {with/that have} ambition

The foregoing generalization appears to extend to other determiners listed in (1), as may be seen from (10)–(16):

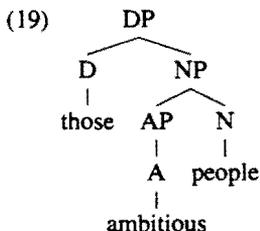
- (10) Many {*intelligent/more intelligent than you} have tried and failed.
- (11) Most *(who are) successful credit their achievement to hard work.
- (12) Few capable *(of doing the work) are willing to move to our area.
- (13) I culled all *(that were) undesirable.
- (14) We hired both *(who appeared) knowledgeable.
- (15) I located two {*worthwhile/worth keeping}.
- (16) Mary's old article on clitics is better than John's {*recent/on anaphora}.

In each case, the acceptable alternative is the more complex attributive modifier, capable of occurring post-nominally in a normal DP featuring a common noun. Furthermore, there is a small class of simple adjectives that may occur post-nominally, and the data in (17) and (18) suggest that these too may be employed even when no common noun is present:

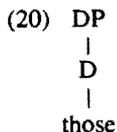
- (17) a. all present
- b. all persons present
- (18) a. those available
- b. those persons available

While the generalization presented here appears fairly robust, and while it seems simple enough, it turns out to be surprisingly difficult to implement under standard assumptions about phrase structure. I shall show below, however, that it follows straightforwardly from an analysis employing Lexical Sharing.

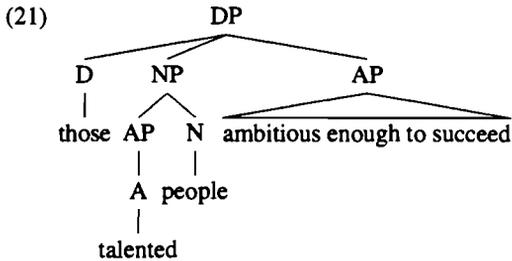
At present, let us set about reviewing some possible conventional analyses of the contrast in (4) and (5). First recall that the DP analysis generally posits structures like (19) for normal nominal structures in which common nouns are present:



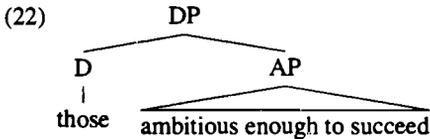
Under the DP hypothesis, the very simplest analysis of cases where *those* doesn't precede a common noun would be to suppose that the determiner heads a DP with no NP, as in (20):



Obviously any NP-internal modifier would then be banned. Thus, the success of this analysis in predicting the contrast in (4) and (5) depends on arguing that pre-nominal modifiers are NP-internal, while post-nominal ones attach to DP, in the manner illustrated in (21):



Only then would one be able to provide an account of such grammatical forms as (5a), by positing a structure like (22):



Positing the structures in (21) and (22) runs counter to standard assumptions about attributive-modifier syntax, though. I will consider two types of conflicting evidence.

The first argument against the structures in (21) and (22) concerns coordination. Note that in (23)–(26) it is possible to coordinate a noun with a post-nominal modifier with a subsequent noun that lacks a post-nominal modifier; in each case the determiner or quantifier may take scope over the whole coordination:

- (23) *The paint peeling off the wall* and broken windows gave the place a desolate look.
- (24) Several *men dressed in western suits* and kimono-clad women emerged from the temple.
- (25) Every *cop on the take* and dishonest public official will be rooted out and prosecuted.
- (26) Any *doll that talks* or other animated toy will be a big hit this Christmas.

These data may be straightforwardly analyzed as NP coordination, if one assumes that post-nominal modifiers attach at the NP level. However, if one maintains that post-nominal modifiers attach at the DP level, then the examples in (23)–(26) would all feature coordination of an NP+XP sequence with an NP. Even in non-constituent coordination, such an aparallelism would be disallowed; witness (27):

- (27) *John kicked a soccer ball to Mary and a football.

Thus, there seems to be little choice but to assume that post-nominal modifiers attach at the NP level.

Another argument against assuming that post-nominal modifiers are daughters of DP comes from certain constructions where one may optionally use an NP in place of a DP. For instance, certain verbs, like *appoint*, *elect*, *name*, and *proclaim*, optionally take NP complements, rather than DPs. Similarly, NPs may be used in place of DPs in appositives. Thus, as illustrated in (28), even superlatives may occur without a determiner in these environments, whereas elsewhere determinerless superlatives are strictly ungrammatical:

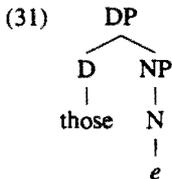
- (28) a. Ed Wood was proclaimed *worst director of all time* by a panel of critics.
 b. Ed Wood, *worst director of all time*, brought us such flops as *Plan 9 from Outer Space*.
 c. **Worst director of all time* was Ed Wood.

Now, notice that NPs in these environments may take post-nominal modifiers, as in (29) and (30):

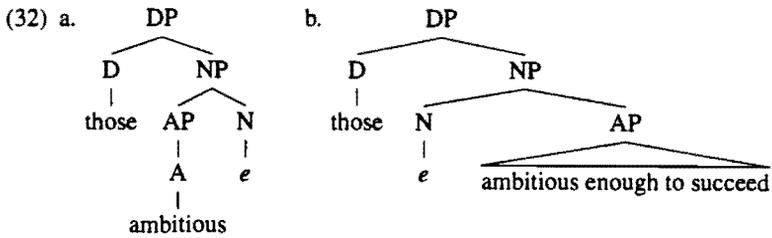
- (29) a. Mary was appointed *vice president in charge of procurement*.
 b. Mary Smith, *vice president in charge of procurement*, declined to comment.
 c. **Vice president in charge of procurement* declined to comment.
 (30) a. Mary was elected *congresswoman representing the state of California*.
 b. Mary Smith, *congresswoman representing the state of California*, declined to comment.
 c. **Congresswoman representing the state of California* declined to comment.

Thus, it would appear that post-nominal modifiers are contained within the NP. As a result of these data and the foregoing coordination facts, it seems clear that one cannot analyze determiners without accompanying common nouns simply as instances of DP lacking an NP daughter.

An alternative to the foregoing analysis would be to posit an NP headed by an empty category, as in (31):

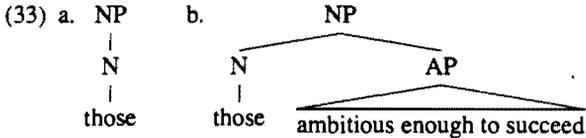


However, this affords no solution, since there is no principled reason to distinguish between (32a) and (32b):

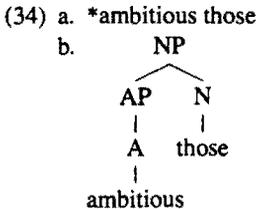


Under this analysis, the choice between (32a) and (32b) would have to be simply stipulated.

Yet another possibility might be to suppose that what appear to be determiners without accompanying common nouns are in fact instances of the category N. One would then have trees approximately like those in (33):

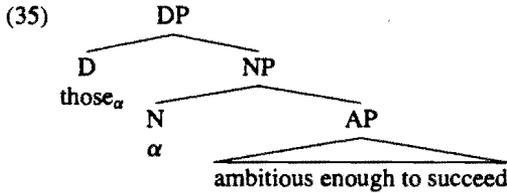


One could, without affecting this argument, embed these NPs inside of DPs whose heads are either inaudible or altogether absent. Note that such an analysis would allow all of the grammatical forms in (5), while ruling out (4), **these ambitious*. In the ungrammatical **these ambitious*, the simple adjective would illicitly follow the head noun. However, this sort of analysis runs into a new problem; there is no principled way to rule out such forms as (34):



In other words, one would have to find some way of ensuring that the noun *those* could have any of the usual post-nominal modifiers but no pre-nominal adjectives. However, there is no obvious reason why a structure like (33b) is to be preferred over one like (34b). Thus, the analysis embodied in (33) can only rule out (34) by means of an ad hoc stipulation.

At this point we have considered and found fault with the three most obvious analyses that could be applied to the problem of predicting the contrast in (4) and (5). Consequently I shall next turn to Lexical Sharing, which solves the above problems by allowing phrase-structure representations like (35):



Recall that Greek variables indicate sameness of lexical tokens; i.e., D and N are associated with the very same word. Now, attributive APs without complements, which *must precede* the head noun, would fall between the N and D in (35), violating the lexical integrity of *those*, which is 'shared' by these two nodes. Thus, one correctly predicts that only attributive modifiers that *may follow* the head noun can co-occur with lone determiners. I shall return to these empirical considerations after providing an axiomatization of Lexical Sharing.

3. ON LEXICAL SHARING

There are few published studies that question the basic assumptions underlying phrase-structure trees.² One senses a prevailing mind-set that accepts these representations with their familiar properties as a given. It is therefore worth stressing that trees are mathematical structures with underlying definitions.³ Moreover, those definitions are susceptible to modification to create mathematical structures with different properties. This work is based on such a redefinition of trees.

In effecting this redefinition of phrase-structure trees, I shall follow the classical axiomatic method, defining primitive individuals, functions, and relations, as well as axioms that determine their behavior. Thus, to begin, a syntactic phrase-structure tree, denoted as in (36), is defined by the nine primitive notions described in (37):⁴

(36) $T = (N, L, \Lambda, \Sigma, l, \lambda, \sigma, D, <)$

(37) *Primitive Notions Defining a Tree*

- | | | |
|-------------|---|---|
| N , | a finite set of phrase-structure nodes, | variables: \dots, w, x, y, z ; |
| L , | a finite set of syntactic-category labels, | variables: a, b, c, d, \dots ; |
| Λ , | a finite set of lexical tokens, | variables: $\alpha, \beta, \delta, \gamma, \dots$; |
| Σ , | a finite set of lexical forms, | variables: $\dots, \varphi, \chi, \psi, \omega$; |
| l , | the label function from N onto L ; | |
| λ , | the lexical-token function from leaf nodes onto Λ ; | |
| σ , | the lexical-form function from Λ onto Σ ; | |
| D , | the 'dominates' relation, a weak partial ordering of N ; | |
| $<$, | the lexical 'precedes' relation, a linear ordering of Λ . | |

The axioms of the theory, which determine the properties of trees, will be stated in a many-sorted, first-order logical language with four groups of sorted variables drawn from the initial and final portions of the Roman and Greek alphabets, as summarized in (37). In what follows, I provide intuitive descriptions of the primitives, followed by formal specifications of the axioms and various auxiliary definitions.

3.1. *Primitive Individuals*

The theory posits four sorts of primitive individuals: *nodes* in N , which model syntactic phrases, *labels* in L , which model syntactic categories, *lexical tokens* in Λ , which model word tokens strung together linearly in an utterance, and *lexical forms* in Σ , which model words viewed as discrete semantic units. As for the distinction between these last two sorts of individuals, note that in the sentence *The boy read the book*, the two instances of *the* are distinct lexical tokens but correspond to the same lexical form. The traditional notion of trees makes do with the first two sorts; lexical tokens are treated as nodes, and lexical forms as labels. Thus, a node labeled N might 'dominate' a node labeled *book*, in the usual sense. However, I wish to introduce Lexical Sharing, which cannot be modeled by traditional domination. Consequently I distinguish *phrasal* entities—i.e., nodes and labels—from *lexical* ones—i.e., lexical tokens and forms. Domination will be a relation strictly among nodes, and the association between nodes and lexical tokens will be handled with a separate function, to be described next.

3.2. *Association Functions*

Labels from L are associated with each node by the function l . Thus, each node has a unique label, which indicates the node's syntactic category. The function λ associates each leaf node with a lexical token in Λ , where 'leaves' are nodes that dominate no other nodes. Each lexical token is in turn associated with a lexical form in Σ by the function σ . It is with λ that I implement Lexical Sharing, which drives the present analysis. Were one to model usual linguistic practice, one would make λ a one-to-one mapping from leaf nodes onto Λ . In contrast, I assume only that λ has the basic properties of a function. Thus, λ associates each leaf node with at most one lexical token; however, it is permissible for λ to relate two or more leaves to the same token. In the latter case, one may say that the lexical token is 'shared' by multiple leaf nodes.

3.3. *Domination*

I shall make the D relation model the familiar properties associated with 'domination' in the literature. These are precisely embodied in graph-theoretic directed trees, so I look to that discipline for my axioms. To begin, I require the 'immediately dominates' relation ID in (38):

(38) **Definition:** *The 'Immediately Dominates' Relation ID*

$$ID(x, y) \leftrightarrow \forall z(D(z, x) \leftrightarrow (D(z, y) \wedge z \neq y))$$

A node x *immediately dominates* a node y precisely when the nodes that dominate x are exactly those that dominate y , less y itself.

The axioms in (39) and (40) echo a standard graph-theoretic definition of directed trees (Thulasiraman & Swamy 1992, p. 106 ff.). Note that domination is described with kinship metaphors: *ancestors* dominate *descendants*, and *mothers* immediately dominate *daughters*.

(39) **Axiom: Rootedness**

$$\exists r \forall x D(r, x)$$

There must be a *root*, i.e., a node which dominates all nodes.

(40) **Axiom: Mother Uniqueness**

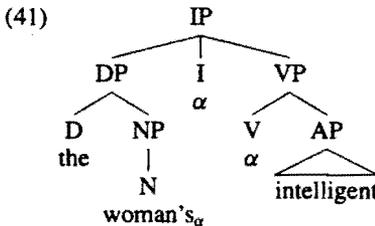
$$\forall x \forall y \forall z ((ID(x, z) \wedge ID(y, z)) \rightarrow x = y)$$

A node may have at most one mother.

3.4. *Left-to-Right Ordering*

I next present a new view of left-to-right ordering based on the $<$ relation, which is a linear ordering on lexical tokens representing the obvious single-file temporal sequence of words in an utterance. To model left-to-right ordering at the phrasal level, I provide the *loose precedence* relation $<_L$, which is defined as an abstract derivative of the lexical ordering $<$.

The tree in (41), featuring a contracted copula, illustrates the need for a novel type of precedence relation at the phrasal level:



If one takes a strict lexicalist approach maintaining that *woman's* must emerge from the lexicon as a ready-made unit, one can still accommodate the usual assumptions about constituency by adopting a Lexical-Sharing approach, as in (41). However, in this structure the subject DP and the VP are not strictly ordered in the usual sense, because it is not the case that every lexical token belonging to the DP precedes every token belonging to the VP. Thus, if one is to be able to make obviously relevant statements like 'the subject DP precedes the VP', one requires a lenient ordering relation which tolerates a moderate overlap of one shared lexical token between phrases. The loose precedence relation is designed to satisfy this need.

Let us now define the loose precedence relation, $<_L$. To effect this definition, it is convenient to introduce one auxiliary notion; let the *Covers* relation be the relative product of D and λ :

(42) **Definition: The Covers Relation**

$$\text{Covers}(x, \alpha) \leftrightarrow \exists y (D(x, y) \wedge \lambda(y) = \alpha)$$

A node x *covers* a lexical token α precisely when x dominates a node that λ maps into α .

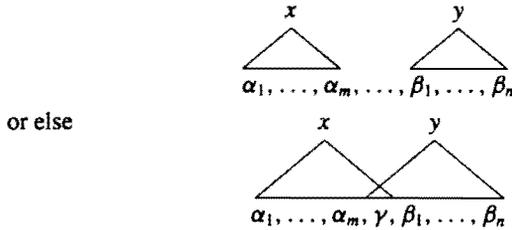
In classical definitions of trees, the *Covers* relation would be subsumed by domination; however, here the new relation is necessitated by the formal distinction between D and λ . Loose precedence may now be defined as in (43):

(43) **Definition:** *The Loose Precedence Relation on Nodes, $<_L$*

$$x <_L y \leftrightarrow \forall \alpha \forall \beta ((Covers(x, \alpha) \wedge Covers(y, \beta)) \rightarrow (\alpha < \beta \vee \alpha = \beta))$$

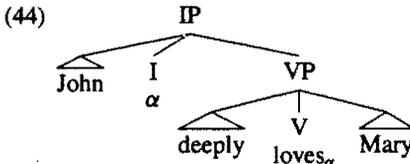
A node x *loosely precedes* a node y precisely when every lexical token covered by x is as far to the 'left' as every token covered by y .

Illustration: Schematically, x loosely precedes y in either of the following configurations, where x covers $\alpha_1, \dots, \alpha_m$, y covers β_1, \dots, β_n , both x and y cover γ , and $\alpha_i < \gamma < \beta_j$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$:



As a result of this definition, for any nodes x and y , if $x <_L y$ holds, then x and y may cover at most one lexical token in common, and that token will be the rightmost one covered by x and the leftmost one covered by y . Now let us reconsider the case of *woman's* in (41); the subject DP loosely precedes the VP, and the fact that both cover the lexical token *woman's* is no longer a problem. Thus, $<_L$ successfully models the intuitive notion of precedence called for above.

It should be noted at this point, however, that not every pair of sister nodes stands in the loose precedence relation. For instance, in (44), the I node covers *loves*, even though the latter is neither the leftmost nor the rightmost token covered by the VP:



In such cases, let us say that the VP *overlies* the I; this is defined in (45):

(45) **Definition:** *The Overlies Relation*

$$Overlies(x, y) \leftrightarrow \forall \alpha (Covers(y, \alpha) \rightarrow Covers(x, \alpha))$$

A node x *overlies* a node y precisely when x covers every lexical token covered by y .

I speculate that given two sister nodes not ordered by $<_L$, one may overlie the other only if the latter is a leaf node. This merits empirical investigation, but for

the moment I will content myself with this assumption and incorporate it into the present theory.

Employing the formal definitions of sister- and leafhood in (46) and (47), one may posit the Configurationality Axiom in (48):

(46) **Definition:** *The Sisterhood Relation*

$$\text{Sisters}(x, y) \leftrightarrow (x \neq y \wedge \exists z(\text{ID}(z, x) \wedge \text{ID}(z, y)))$$

Nodes are *sisters* precisely when they are distinct yet have the same mother.

(47) **Definition:** *The Leafhood Property*

$$\text{Leaf}(x) \leftrightarrow \forall y(D(x, y) \rightarrow x = y)$$

A *leaf* is a node with no distinct descendants; i.e., it dominates only itself.

(48) **Axiom:** *Configurationality*

$$\forall x \forall y (\text{Sisters}(x, y) \rightarrow (x <_L y \vee y <_L x \vee (\text{Leaf}(x) \wedge \text{Overlies}(y, x)) \vee (\text{Leaf}(y) \wedge \text{Overlies}(x, y))))$$

Sisters either stand in the loose precedence relation or else one is a leaf which the other overlies.

As a result of the Configurationality Axiom in (48), the phrase-structure trees defined here generally have the same left-to-right ordering properties as traditional constituent structures. Only when Lexical Sharing is present do certain unconventional precedence relationships arise among nodes, and these are contained within limited regions of the tree. This issue is discussed in greater detail with formal proofs in Wescoat 1996.

4. THE PARADOX RESOLVED

With the foregoing axiomatization in place, one may return to the empirical issues raised in section 2. The problem outlined there is obviously the product of an ordering paradox, so let us begin by using the newly defined loose precedence relation to impose the ordering constraints in (49) on the elements inside of DP:

(49) *Ordering Principles within the DP*

- a. D must loosely precede NP.
- b. AP must loosely precede N, if the former has no complements.
- c. Otherwise, N must loosely precede all XP.

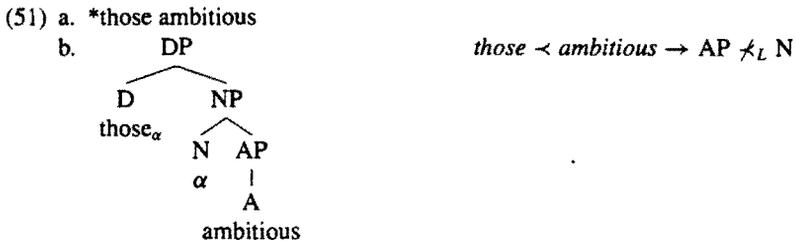
I have presented the statements in (49) as axioms for specificity, but it would obviously be more interesting if they could be shown to follow as theorems from some more basic principles. Although I shall not pursue the matter here, the work of Sadler and Arnold (1994) shows promise of providing the foundation for such an approach. In any case, (50) will suffice to derive all the relevant results for this section. Thus, let us set about reconsidering various crucial data from section 2 along with their Lexical-Sharing analyses.

First of all, consider the basic structure in (50), which contains no modifiers:



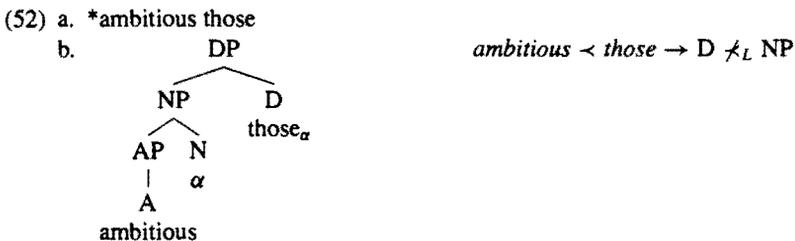
Here, D and NP both cover the single lexical token *those*, so there is no token covered by NP that precedes any token covered by D. As a result, it follows from the definition of loose precedence that D loosely precedes NP. Thus, the only relevant ordering principle, (49a), is satisfied, and the tree in (50) is admitted.

Next let us consider the ungrammatical string in (51):



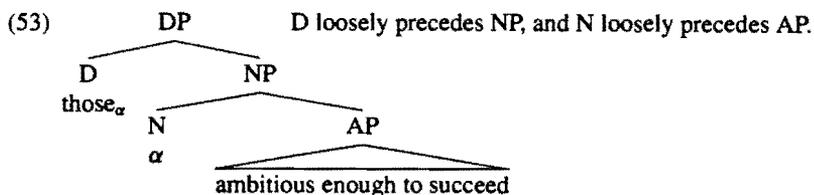
As indicated in the implication to the right of the tree in (51), the fact that *those* lexically precedes *ambitious* means that at least one lexical token covered by N precedes a token covered by AP, whence it follows that AP does not loosely precede N. Yet, this particular AP has no complements; thus, by (49b) it should loosely precede N, whence the infelicity of (51). Note that the foregoing conclusion is entailed solely by the fact that *those* lexically precedes *ambitious*; thus, though one might have an intuitive impulse to try to alter the outcome by redrawing the tree in (51) with the N to the right of the AP, the facts regarding lexical precedence would remain unchanged and would lead to the exact same result.

Similar reasoning holds in (52), where *ambitious* precedes *those*:



Since *ambitious* lexically precedes *those*, D does not loosely precede NP. As this violates (49a), the tree in (52) is ruled out.

Finally, we have the tree in (53), which represents all of the structures in which a post-nominal type modifier occurs in conjunction with *those*:



Here, D covers only the leftmost lexical token, *those*, while NP covers all five tokens. Thus, obviously D loosely precedes NP, satisfying (49a). Though I have not given a full tree, one may take it for granted that *those* will not be lexically shared by any descendant of AP. As a result, *those*, the single lexical token covered by N, precedes every token covered by AP, whence it follows that N loosely precedes AP, in conformity with ordering principle (49c). Thus, the tree in (53) is not ruled out by the ordering principles, the way (51) and (52) were. In sum, it appears that the Lexical-Sharing analysis along with the very basic ordering principles stated in (49) suffices to capture the contrast in (4) and (5), while also ruling out (34a).

Interestingly, the effects discerned in the foregoing data are also found in conjunction with other lexical classes associated with the category D. For instance, pronouns have been identified as determiners on the basis of data like (54)–(56):

(54) *We teachers* are doing the impossible.

(55) *You students* need to buckle down.

(56) *Them things*'re good eatin'.

Pronouns appear, furthermore, to admit attributive modifiers according to the same pattern as was found in (4) and (5), as seen in (57)–(61), where only the simple adjective in (57) is deemed unacceptable:

(57) **They* make mistakes that we experienced can only deplore.

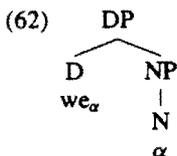
(58) *They* make mistakes that we more experienced in the ways of computers can only deplore.

(59) *We* in the diplomatic corps have experience in these matters.

(60) *You* who have everything cannot understand our plight.

(61) *He* that laughs last laughs best.

To account for these facts, we can assign pronouns the structure in (62), which is analogous to the one posited above for normal determiners unaccompanied by common nouns:



Wh-words are also susceptible to an analysis in which they assume the category D, as (63) suggests:

(63) What books should I read?

Once again, one finds the same pattern of acceptability with attributive modifiers; this is demonstrated for *what* in (64)–(67), and for *who* in (68)–(71):

(64) *What important could he have up his sleeve?

(65) What important enough to arouse the boss's interest could he have up his sleeve?

(66) What in the world could he have up his sleeve?

(67) What that I could possibly know about would you be interested in?

(68) *Who smart would take the job on?

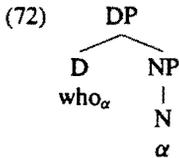
(69) Who smart enough to do the job would be dumb enough to take it on?

(70) Who in your office is capable of completing this project?

(71) Who that you know is likely to come to the concert?

(McCawley 1988, p. 369)

Thus, we might posit the structure in (72) for *wh*-words:



Now, in contrast with *what* in (63), the *wh*-word *who* never functions as a determiner before a common noun, as suggested by (73):

(73) *Who people will come?

This fact can be readily accommodated by stipulating that *who* is always lexically shared by a D and an N, as in (72), while such sharing is optional for *what*.

Thus, the foregoing analysis seems to be of general utility in analyzing a broad range of determiner constructions in English.

5. CONCLUSION

In conclusion the Lexical-Sharing analysis appears to provide a very straightforward account of some subtle facts about the structure of nominal expressions in English. To the best of my knowledge, the types of facts discussed here have not been dealt with in recent studies of English syntax, so it is difficult to compare the present approach to other theories. However, one can say that among constituent analyses available under the traditional conception of phrase structure, there do not appear to be any obvious accounts capable of handling the facts. Furthermore, among the notions which recent theories introduce to supplement the power of their grammars, such as movement transformations, feature checking, or unification, there doesn't

appear to be any tool that offers an obvious solution to the problems posed here. Thus, I conclude that Lexical Sharing provides an essential mechanism for detailed syntactic analysis.

NOTES

*I wish to express my gratitude to Ernest W. Adams, whose kind guidance has always been indispensable to the formal development of the notion of Lexical Sharing employed here. I am also greatly indebted to Christopher Culy, Jason Merchant, Geoffrey Pullum, Peter Sells, Thomas Wasow, and Arnold Zwicky for insightful discussion of this material. Of course, responsibility for the remaining deficiencies in this work rests solely with me.

¹Of course 'people deletion' constructions like *the ambitious* are well formed. Such forms are susceptible to an analysis similar to that put forward in this study, though I shall not discuss the matter further here.

²McCawley 1982 and Blevins 1994 are notable exceptions to this generalization.

³Partee, ter Meulen, and Wall (1990, p. 439 ff.) axiomatize traditional trees.

⁴Here I offer fewer primitives than in Wescoat 1994 and 1995. This is the minimal set needed to implement Lexical Sharing; my previous papers required additional relations to model \bar{X} -theory along with argument- and adjuncthood.

REFERENCES

- Abney, Steven Paul: 1987, *The English Noun Phrase in Its Sentential Aspect*, Ph.D. dissertation, MIT, Cambridge, MA.
- Blevins, James P.: 1994, 'Derived Constituent Order in Unbounded Dependency Constructions', *Journal of Linguistics* 30, 349–409.
- Fukui, Naoki: 1986, 'SPEC, Agreement, and the Comparative Syntax of English and Japanese', *NELS* 17, 193–210.
- McCawley, James D.: 1982, 'Parentheticals and Discontinuous Constituent Structure', *Linguistic Inquiry* 13, 91–106.
- : 1988, *The Syntactic Phenomena of English*, University of Chicago Press, Chicago.
- Partee, Barbara H., Alice ter Meulen, and Robert E. Wall: 1990, *Mathematical Methods in Linguistics*, Kluwer, Dordrecht.
- Sadler, Louisa, and Douglas J. Arnold: 1994, 'Prenominal Adjectives and the Phrasal/Lexical Distinction', *Journal of Linguistics* 30, 187–226.
- Thulasiraman, K. and M. N. S. Swamy: 1992, *Graphs: Theory and Algorithms*, John Wiley & Sons, New York.
- Wescoat, Michael T.: 1994, 'Phrase Structure, Lexical Sharing, Partial Ordering, and the English Gerund', *BLS* 20, 587–598.
- : 1995, 'Lexical Sharing in a Monostratal Syntax of English Wh-Constructions', *CLS* 31, 490–507.
- : 1996, *Hindi and the Typology of Noun Incorporation: An Analysis with Lexical Sharing*, Osaka, Japan: Osaka University, MS.