Linguistik Aktuell Linguistics Today

# Asymmetry in Grammar

## Volume 2: Morphology, phonology, acquisition

*Edited by*
Anna Maria Di Sciullo

Asymmetry in Grammar

# Linguistik Aktuell/Linguistics Today

*Linguistik Aktuell/Linguistics Today* (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.




**Volume 58**

Asymmetry in Grammar: Volume 2: Morphology, phonology, acquisition
Edited by Anna Maria Di Sciullo

# Asymmetry in Grammar

## Volume 2: Morphology, phonology, acquisition

Edited by Anna Maria Di Sciullo
University of Quebec at Montreal

**John Benjamins Publishing Company**
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

**Library of Congress Cataloging-in-Publication Data**

Asymmetry in Grammar : Volume 2: Morphology, phonology, acquisition / Edited by Anna Maria Di Sciullo.
p. cm. (Linguistik Aktuell/Linguistics Today, ISSN 0166–0829 ; v. 58)
Includes bibliographical references and indexes.
Contents: v. 1. Syntax and semantics -- v. 2. Morphology, phonology, acquisition.
1. Asymmetry (Linguistics) I. Di Sciullo, Anne-Maria, 1951- II. Linguistik aktuell ; Bd. 57-58.

P299.A85 A845 2002
415-dc21 2002033272
ISBN 90 272 2778 0 (v.1 Eur.) / 1 58811 306 X (v.1 US) (Hb; alk. paper)
ISBN 90 272 2779 9 (v.2 Eur.) / 1 58811 307 8 (v.2 US) (Hb; alk. paper)
ISBN 90 272 2783 7 (set Eur.) / 1 58811 320 5 (set US)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

# Table of contents

**Acquisition**

# Asymmetry in grammar

## Morphology, phonology, acquisition

Anna Maria Di Sciullo

### Introduction

The Set theoretical notion of asymmetry and symmetry are used to define the properties of the operations of the grammar. In Chomsky (2000), set-Merge, which derives the Head complement relation is symmetric, whereas pair-Merge, the adjunction operation, is asymmetric. In Kayne (1994), the linear order of terminals in a syntactic tree is a function of the complete set of ordered pairs of pre-terminals for which asymmetric c-command holds.

While research has been mainly focussed on the properties of syntactic asymmetries, Chomsky (1998, 1995, 2001), Kayne (1984, 1994, 2001), there have also been works on asymmetry in morphology, phonology, as well as in acquisition.

In morphology, the rules of word formation derive asymmetric, couched in terms of the notion of Head of a Word (Selkirk 1982; Williams 1981) and Relativized Head of a Word (Di Sciullo & Williams 1987). Asymmetry is also part of the Adjunct Identification Condition (Di Sciullo 1997) proposed to account for the restrictions on the adjunct-head relations, viz. adjuncts identify unspecified features of heads. An empirical consequence of these proposals is that a prefix cannot be the categorial head of a word. Moreover, asymmetries in composition have also been observed. For example, subjects and prepositional complements are rare or unacceptable in deverbal compounds. This fact has been proposed to follow from construction specific principles, such as the First Sister Principle (Roeper & Siegel 1978). Morphological asymmetries in compounds are also discussed in Di Sciullo and Williams 1987, Kayne 1994, Lieber 1992, Di Sciullo 1996 and in Roeper and Keyser 1997.

The hypothesis that morphological expressions include a head as well as dependent projections, as it is the case in syntactic expressions, allows for a unified treatment of the different expressions derived by the grammar. The question arises whether there should be any difference at all between syntax and morphology. This question is still subject to debate, as can be seen with the different viewpoints embodied by Distributed Morphology (Halle & Marantz 1993) and Amourphous Morphology (Anderson 1992). Baker (1988) proposed to extend the set of transformations in order to include Head Movement. The Head movement constraint (Travis 1984) was proposed to account for the restrictions on the movement of heads. However, Head movement has been shown to be problematic in more than one way (Di Sciullo & Williams 1987; Brody 2001). Koopman and Szabolcsi (1998), Mahajan (2000) have proposed that Head movement is not found in syntax. The theoretical reason is that Head Movement does not satisfy the core properties of syntactic derivations, the Extension Condition and the Inclusiveness Condition (Chomsky 1995). This reopens the question of the proper treatment of displacement in both syntax and morphology, as well as the proper derivation of both syntactic and morphological asymmetries.

Asymmetry in phonology has been discussed with respect to feature geometry, head dependencies, and more recently with respect to the PF linearization. Work on the distributional inequalities between positions in the prosodic hierarchy reflecting asymmetries in the prosodic licensing relations are important in this area (Harris 1994, 1997; Huslt & Ritter 1999).

A source of asymmetry in phonology is the presence of head-dependency relations (Dresher & van der Hulst 1998; Dresher & Rice 1993; Hulst 1984, 1999, 2000; Piggot 1999, 2000; Rice 1992; Rice & Avery 1993). There is an interaction between the asymmetry that results from head-dependency relations and the symmetry that lies in the fact that phonological oppositions are typically polar paradigmatic opposites (like oral–nasal, stop–continuant, front–back, etc.). Head-dependency relations are also at play in the syntagmatic domain (syllable structure and beyond). The question arises whether the notion of asymmetry and the notion of head-dependency can be identified as identical. The question also arises whether the phonological asymmetries are isomorphic to the syntactic and the morphological asymmetries. Papers from this collection suggest that this is not the case. Raimy (2000, in this volume) argues that the linearization process in phonology is similar to but distinct from the syntactic LCA (Kayne 1994) in fundamental ways. Similarities between the two are that linearization ensures asymmetry in representations at an interface, and that both linearization processes appear to be language universals. The differences between the two linearization processes are more pronounced. The LCA creates default precedence relationships between morpho-syntactic units that previously did not have any precedence relationships, whereas phonological linearization ensures that a precedence structure is interpretable at the phonetic interface. The LCA eliminates symmetric relations (as created by movement) by 'deleting' parts of a syntactic chain whereas phonological linearization repeats segmental material (without identity between the copies). Because the linearization processes in syntax and phonology are different, we can expect non-trivial interactions between the two processes.

If phonological asymmetry exists, as well as syntactic and morphological asymmetries, the question that comes to mind is whether asymmetry is given by Universal Grammar. It might be the case that asymmetry is part of the initial state of the Language Faculty, enabling human beings to develop the grammar of the language to which they are exposed, to interpret and to quickly generate the expressions of this language in a short period of time.

This collection of paper contributes to the view that asymmetry is basic in Grammar and thus is a property of grammatical relations across the board. The following paragraphs summarize the main contributions of each paper.

## Morphology

In 'Morphological Relations in Asymmetry Theory', Anna Maria Di Sciullo approaches the properties of morphological relations with Asymmetry Theory. According to this theory, morphological objects cannot reduce to either non branching or binary branching configurations. Even though such configurations have been proposed in the literature for the structure of wh-expressions, she lays out theoretical as well as empirical motivation to show that these expressions, as it is the case more generally for morphological expressions, are minimally articulated in terms of shells, that is, double-layered asymmetric structures. She shows that the specifier-head and the head-complement relation must be projected in the higher and in the lower layer of the shell, in order for the feature structure of the functional constructs to be properly deployed. She brings support to show that the structuring of functional features in morphological objects is not isomorphic to their syntactic structuring, even though morphological and syntactic objects are built on the basis of asymmetric relations. Predictions are made with respect to the restrictions on composition and ordering of functional features in wh-words and subordination conjunctions in English and Romance, more specifically Italian, French, and Romanian.

In 'Asymmetries in Prefixation vs. Compounding: The Case of Greek Preverbs', Angela Ralli deals with the role of three Greek preverbs, *ksana, kse-*, and *para-* in verb formation. By examining their characteristics on semantic, structural and phonological grounds, she proposes a morphological approach that allows for a unified treatment of the three preverbs and accounts for both their differences and similarities. She shows that a distinction into internal and external preverbs, or a division in prefixes and non-prefixes, is not sufficient to take into consideration their peculiar behavior. She argues that the combination of the preverbs with the root verbs is an instance of a word-formation process that is handled within a morphological module of grammar and refers to either derivational prefixation or compounding. Following Di Sciullo 1997, she considers that the combination of the preverbs with the verbal bases occurs at different adjunction sites in grammar. However, she proposes that these adjunction sites can be expressed in terms of morphological categories, i.e., as adjunction to stems or adjunction to words. She shows that a difference in adjunction sites within morphology can explain the asymmetrical character of certain preverbs that behave like prefixes but share more similarities with word-like preverbs, than with prefix-type ones.

In 'Asymmetry and the Typology of Formal Objects', Réjean Canac-Marquis discusses an approach based on the fundamental notion of Asymmetry in the sense of Di Sciullo (1999). According to Di Sciullo's hypothesis, all grammatical relations must be asymmetric. He explores the consequences of this hypothesis for the domain of formal objects. He proposes that all formal objects, all projections, must be interpretable in terms of the fundamental asymmetric relations of immediate dominance and dominance. Hence Xmax and XNmax are not primitives but are contextually interpreted in terms of local (immediate) dominance and non-local dominance, asymmetric relations. His proposal predicts a typology of formal objects and advantageously subsumes stipulative aspects of Chomsky's system, e.g. derives the effects of the Chain Uniformity Condition. The typology also does away with exceptions in Chomsky's system, namely that head-adjunction should be excluded in principle, but is permitted in practice. He maintains that head-adjunction, not head substitution, is excluded in principle in syntax, explores the consequence of this proposal for the Mirror Principle effects, and argues that the absence of head-adjunction in syntax is the fundamental explanation behind the Mirror Generalization, not its source as claimed in Baker (1988, 1998).

In his paper 'Analytic/Synthetic Variation as Language Specific Asymmetry', Abdelkader Fassi Fehri begins with some typological generalizations across languages which end in a formal claim on the "strategies" that languages use:

A formal relation (R) is either synthetic or analytic in a language. To instantiate this claim, he examines how the perfect and past tenses and passive voice are expressed in various languages (mainly French and English), and how they contrast with their Arabic equivalents. His other main aim is to show that Formal Complexity drives the change from synthetic to analytic expressions. He also discusses various structures in French and Arabic where there is a "split" to be found (e.g., subject/object agreement in French, or the avoir/être split). He also mentions split agreement in Latin perfect tenses. He touches upon Modern Greek tense/voice combinations and even the passive in Albanian.

## Phonology

In 'Asymmetry and Linearization in Phonology', Eric Raimy presents the relevant aspects of asymmetry and symmetry in phonological representations and compares these traits with the asymmetric characteristics of syntax (Kayne 1994). He shows that the traditional way precedence is encoded encounters limitations. He presents a new model where precedence takes the form of asymmetric and symmetric relations, instead of assuming that precedence is derived from left-to-right graphic ordering. Once the revised model of precedence is adopted, more complex and previously unconsidered phonological representations can be considered. He proposes that phono-logical representations contain loops to produce reduplicated forms due to a linearization process that eliminates symmetric relations (assuming that *precedence* is transitive) through repetition. Evidence for symmetrical looping precedence structures at some point in a phonological derivation is provided by backcopying phenomena in reduplication and the interaction between infixing and phonological rules.

In 'Levels, Constraints, and Heads', Harry van der Hulst and Nancy Ritter offer a general outline of a two-level (i.e. lexical, post-lexical), principles-and-parameters approach to phonology that, effectively, is a constraint-based system because it does not use extrinsic ordering of phonological operations at either level. They discuss the derivational side of their approach, and present their commitment to a theory of phonological representations that attributes a fundamental role to the asymmetric relation of head-dependency. They propose a refinement of parameter theory that, in fact, extends the role of the head-dependency asymmetry into the domain of parameter setting. While this proposal is primarily made in the spirit of developing their own approach to phonology, it can, at the same time, be seen as an answer to certain objections that have been raised against parameter theory by proponents of Optimality

Theory (OT). Implicit in this latter remark is the point that they have not embraced the ideas of OT. In comparing their own constraint-based approach to OT, they examine and criticize the OT-device of constraint ranking.

In 'Obstruent Neutrality in Nasal Harmony', Glyne Piggott considers the generalization that segments in nasal harmony systems are either subject to nasalization or are neutral to the process. He provides a principled explanation for the neutrality of obstruent stops, relating the phenomenon to the robust observation that these are the only segments represented in every phonological system. The core of the analysis is the postulation of an IDENTITY principle, prohibiting changes to underlying obstruent stops. This constraint is present in all languages and cannot be superseded by nasal harmony demands. Other IDENTITY constraints blocking changes to other segment-types are grammatical options and therefore not always in effect. Significantly, underlying obstruent stops are sometimes modified, thereby showing that the Obstruent Identity principle is not an inviolable constraint. Nevertheless, violations are never random. The paper claims than the enforcement of obstruent neutrality in some constructions but not others cannot be explained in the currently dominant paradigm of Optimality Theory. Precedence relations between constraints cannot be arbitrary. An adequate explanation of obstruent stop neutrality must postulate that phonological systems are governed by certain universal principles in combination with certain language specific choices.

In his paper 'Towards a Theory of Fundamental Phonological Relations', Charles Reiss develops the formal account of identity references that Odden (1988) demonstrates is necessary for phonological theory. The paper demonstrates the following: 1. The phonological component must be able to express conditions of identity and nonidentity. 2. The computation of non-identity requires the equivalent of the existential quantifier and of variables ranging over the set of features, since non-identity can be based on any arbitrary feature: a segment *a* is not identical to a segment *b* iff there is any feature F for which *a* and *b* have different values. 3. Autosegmental representation, including Feature Geometry (FG) cannot model non-identity, since geometries do not make use of variables and quantification, so FG is insufficiently powerful to generate what a Feature Algebra (FA) model, with indices and variables, can capture. Such relatively simple cases are critical to understand the representational and computational apparatus of UG. The representational power of FG is insufficient, and since its power is subsumed by that of the FA approach with variables and indices, FG must be abandoned. Phonological UG needs at least enough computational power to express identity and non-identity conditions: this can be achieved by the use of quantifiers. The paper concludes with an examination of how these conditions can be accommodated in a theory requiring that all grammatical relations be asymmetrical.

In 'Contrast and Asymmetries in Inventories', Elan Dresher focuses on a source of asymmetry in phonology that derives from the contrastive function of features. He assumes that feature specifications are, at least to some degree, driven by considerations of contrast. A segment must be sufficiently specified so that it is differentiated from other segments in the inventory: the greater the number of contrasts, the more specifications will be required. Specifications that are required to draw a contrast between segments are contrastive with respect to those segments; specifications that are not contrastive are redundant. He argues that in any inventory characterized by more than one distinctive feature, it is necessary to specify the relative scope of each feature. The different featural scope relations correspond to two ways of ordering the contrastive features : in the first scenario, [voiced] >[nasal], while in the second [nasal] > [voiced]. Different contrastive hierarchies create different asymmetries in inventories, and thus different expectations as to how the system will pattern. It is an empirical question whether the contrastive hierarchy is universal, or allows for cross-language variation. He illustrates the workings of the contrastive hierarchy in some detail with respect to synchronic and diachronic patterns in the vowel systems of the Manchu languages.

## Acquisition

In 'The Acquisition of Constructions with Reflexive Clitics in Polish', Maria Louisa Rivero and Magdalena Goledzinowska discuss the acquisition of constructions with reflexive clitics in Polish on the basis of the Childs database and the results of the testing of five children. These are: the reflexive, the reciprocal, the inchoative, the inherent uses and impersonal uses where the reflexive clitic corresponds to "one" (e.g. One opens it this way). The authors support the Maturation hypothesis (Babyonyshev et al., and Borer and Wexler (NLLT 1992). They establish three stages in the process of the acquisition of reflexive clitics: during the first stage, children omit the reflexive clitic in all types of constructions. During the second stage, reflexive clitics occur in intransitive constructions of the unergative and unaccusative types (which contain one argument). And during the third stage, reflexives start appearing in constructions with two arguments, i.e., reciprocals and transitive impersonals.

In 'A Sub Grammar Approach to Language Acquisition', David Lebeaux presents an alternative to the Parameter-setting model of language learning.

He argues that i) the language acquisition process is monotonic, ii) that it starts out with smaller structures, each associated with a licensing relation, iii) that the licensing relations are modeled by generalized transformations, and iv) that as the child grows, the grammar grows from smaller to large, as new generalized transformations, modeling different licensing relations, are added to the grammar. This model of language acquisition is consonant with the hypothesis that the elemental properties of grammar at the initial state are relations and not singular features or categories and that the grammar grows on the basis of the development of elemental relations onto more complex ones. This paper brings developmental evidence for a model of grammar based on relations and not on unstructured elements.

This collection of papers came out from a conference on Asymmetry in Grammar held at the Université du Québec à Montréal in May 2001. I would like to thank the participants to the Conference as well as the members of the Asymmetry Project who helped in the organization and in the process of editing this volume. You will find all their names on our web site: www.asymmetryproject.uqam.ca

Finally, I would also like to thank the Social Sciences and Research Council of Canada for making this Conference possible, and more generally for the financial support to the Major Collaborative Research Project on Asymmetries in Natural Languages and their Treatment by the Performance Systems (Grant no. 412-97-0016).

# Morphological relations in asymmetry theory[1]

Anna Maria Di Sciullo
Université du Québec à Montréal

## 1. Purpose

My purpose is to provide further evidence showing that morphological relations are asymmetric. By morphological relations I mean relations between pair of nodes in M(morphological) trees. I take the set theoretical notion of asymmetry to be basic and the sister-contain relation, in the sense of Chomsky (1998), to qualify as asymmetric. I consider the articulation of features in functional constructs, such as +/– wh features in wh- words, complementizers, and subordination conjunctions. I show that what I call Head-Asymmetry holds within these constructs, as it does more generally in what I call M(orphological)-Shell. As +/– wh constructs are operators, I also refer to their structure in terms of the Op(erator)-Shell.

The derivation of Morphological Shells is forced by conceptual necessity in Asymmetry Theory (Di Sciullo 1999b, 2000 and forthcoming), where asymmetry must be obtained as early as possible for convergence and legibility. M-shells are derived on the basis of two core operations: Shift and Link. These operations are the generalized forms of the Minimalist Merge and Move operations (Chomsky 1995, 1998, 1999). Shift derives complex configurations on the basis of elementary ones; Link relates parts of already Shifted configurations. Given two configurations α, β, Shift derives the configuration δ, projected from either the head of α or the head of β, and Link relates a sub-structure of α to a sub-structure of β under certain conditions. The operations and the conditions apply differently in separate, even though related, planes of the computational space. Thus, I distinguish the M(orphological) from the S(yntactic) Shift and Link operations.

Asymmetry Theory is compatible with the Minimalist Program (Chomsky 1995, 1998, 1999 and related works), as it embodies the Minimalist architecture. To a certain extent, Asymmetry Theory is also compatible with the Antisymmetry framework (Kayne 1994; Moro 2001), as asymmetric c-command (sister-contain) is taken to be a central relation. It is also compatible with Distributed Morphology (Hale & Maratz 1993; Embick 2000), as morphology and syntax both rely on asymmetric relations. It differs from these frameworks with respect to the relative autonomy of morphology and syntax, as well as with respect to the coverage of asymmetry in the grammar.

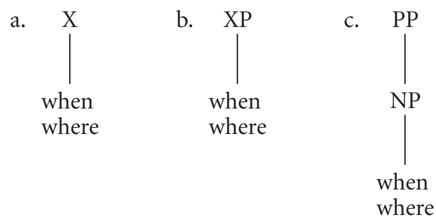
This paper further supports the hypothesis that asymmetry is basic in grammar.[2] It also provides further evidence to show that morphological asymmetries are not isomorphic to syntactic asymmetries.[3] In the first section, I consider the properties of morphological relations. In the second section, I provide evidence that Head Asymmetry holds in functional constructs. As expected, the properties of these constructs are not isomorphic to the properties of functional projections in syntactic structures. In particular, they differ with respect to the articulation of the operator-variable relation, as well as with respect to the variable-restrictor relation. It will become clear that functional Shells are more restricted than the syntactic projection of functional heads. The empirical evidence comes from English, as well as from Romance languages, Italian, French, and Romanian.

## 2. Morphological relations

Morphological objects are structured sets of relations. Given Asymmetry Theory, these relations are asymmetric.[4] I provide evidence to show that (1) holds in functional constructs, as it does more generally in morphological objects. Head Asymmetry is not an independent condition, but it is a consequence of the operations of the grammar. In the case of morphological objects, it falls out from M-Shift directly.

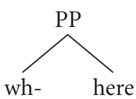
(1) **Head Asymmetry**
A head is asymmetrically related to another category.

Head Asymmetry has consequences for the properties of morphological relations, as it has been established that a morphological expression includes at least one head (Williams 1981; Selkirk 1982; Di Sciullo & Williams 1987). (1) excludes non-branching projections from the set of possible M-relations. Considering +wh-words, such as *where* and *when*, (1) excludes structural descriptions such as (2a) and (2b). However, wh-words have been thought of as elements with no internal structure. This is the case in Larson (1987: 246), where *when* and *where* are analyzed as in (2c).

(2)


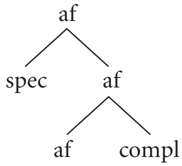

There are reasons to exclude dominance-only relations from the set of possible M-relations. In Asymmetry Theory, a structural relation is not legible if it is not asymmetric. The relations in (2) are not. By definition, dominance, either immediate or non-immediate, is an antisymmetric relation, as it is generally assumed that each node in a tree dominates itself (Wall 1972: 145; Partee, Ter Meuler & Wall 1990: 440).[5]

Head Asymmetry also excludes M-trees such as (3), where the wh- affix and the adverbial PP *here* are sisters, as sisterhood is a symmetric relation.[6]

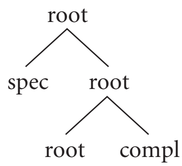
(3)



The sisterhood relation has been assumed to be crucial in the derivation of morphological objects. This is the case for Roeper and Siegel's (1978) First Sister Principle, proposed to account for the restrictions on the derivation of compounds. This is also the case for morphological sub-categorization (Lieber 1980; Grimshaw 1990; Borer 1991; Law 1997), proposed to account for the categorial restrictions on the combination of derivational affixes and roots. In Asymmetry Theory, there is no well-formed tree where a head is not in a sister-contain relation with another category. A morphological head, an affix, does not select for an element it is the sister of, but rather for an element it sister-contains. I have shown elsewhere (Di Sciullo 1996) that a derivational affix projects an asymmetric structure, where the affix itself is part of a head-complement as well as a specifier-head relation, as depicted in (4). Moreover, the affix merges with a root which is also part of a specifier-head and a head-compl relation, as depicted in (5). The merger yields the configuration in (6).

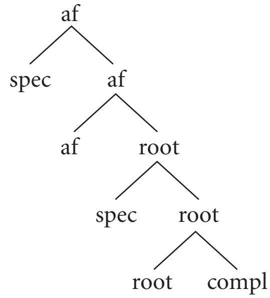
(4)




(5)



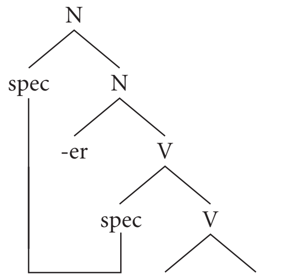

(6)



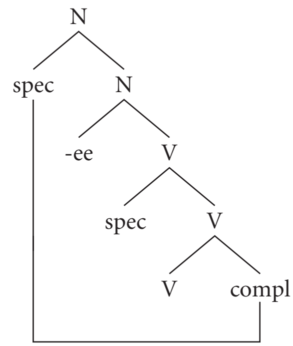

The representations above satisfy Head Asymmetry, as every head they include is asymmetrically related to a spec position. The empirical motivation for (6) is based on asymmetric A(rgument) structure selection. I have shown in Di Sciullo (1996) that this selection is independent of the categorial properties of both the affix and the root. Whether the affix is a nominal, an adjectival, or a verbal category, it projects an asymmetric spec-head-compl configuration and it merges with a root, according to the root's spec and compl A-features. This correctly predicts that certain gaps occur in derivational morphology. Thus, in Asymmetry Theory, a derivational affix does not merge with a category, it merges with a configuration. I refer to this morphological merger as M-Shift. This operation basically combines two trees, T1 and T2, by substituting T2 to the complement position of T1.

The spec and compl positions of a shifted morphological configuration are subject to M-Link. Basically, this operation relates the higher spec to a lower spec or compl, only when the higher spec position differs minimally in feature specification from the lower spec or compl. Thus, taking A-features to be the relevant features here, in a shifted configuration, such as the case of -*er* deverbal nouns, the spec of -*er* is Linked to the spec of the verbal projection, as in (7). In the case of -*ee* deverbal nouns, the spec of -*ee* is linked to the lower compl, as in (8). Likewise, in -*ive* deverbal adjectives, the spec of -*ive* is Linked to the lower spec, as in (9). In the case of -*able* deverbal adjective, the spec of -*able* is linked to the lower compl, as in (10). Given a configurational representation of A-structure (Hale & Keyser 1993), M-Link accounts for the A-structure restrictions on the combination of affixes and roots. Given (7), -er may Shift with an unergative verb but not with an unaccusative verb (e.g. *a sleeper* / * *an arriver).* Given (8), the nominal affix -ee may Shift with a transitive but not with an unergative verb (e.g. *an employee* / **a shinee*). Given (9), the adjectival affix -*ive* may Shift with a transitive verb but not an unaccusative verb (e.g. *an impressive* / * *an arrivative gesture*). Given (10), the adjectival affix -*able* may Shift with a transitive but not with an unergative verb (*an affordable* / * *a sleepable chair*). The regularity of these configurational restrictions reduces the idiosyncratic nature of derivational morphology. I show in sections 3 and 4 that certain restrictions on the formation of functional constructs also follow from Asymmetry Theory.

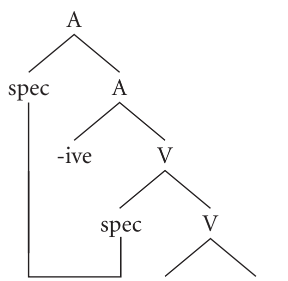
(7)



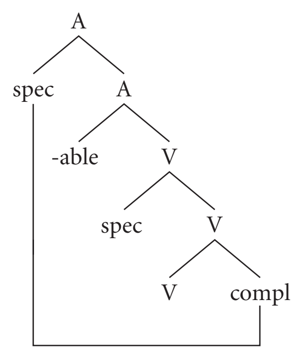

(8)




(9)



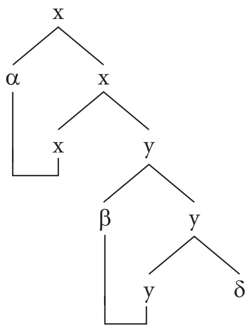

(10)




M-Link performs covert feature attraction/movement. This operation is distinct from the operation Move of Chomsky (1995). Move can be seen as a complex operation including the sub-operations Copy and Merge (see Collins 1997). In its Attract/Move/Delete variant, Move also includes a Deletion operation. By contrast, M-Link performs only one operation. Furthermore, M-Link does not relate two +A or two –A elements, as it is the case in syntax, where when an anaphor is bound by an antecedent, both are +A. Moreover, there is a sharp difference between the morphological spell-out (PF visibility) of the +A and the –A features in the Syntactic and in the Morphological planes. These features are generally spelled out in the S-plane, whereas they are generally not spelled out in the M-plane. In combination, both operations, Shift and Link, allow for a proper derivation of the legible LF feature structure in each plane.

In the next section, I show that asymmetric relations are part of the derivation of functional constructs. I will concentrate on the spec-head relations in these constructs.

(11)




Before doing so, let me summarize the relevant differences between syntactic and morphological relations. Firstly, while morphological shifting is configurational, syntactic shifting (merger) is categorial. In effect, the merger of a syntactic head with a complement is not sensitive to the argument feature of that complement. Secondly, the features in spec and compl in a morphological configuration are not generally legible at PF, whereas these features generally are in syntactic configurations. Thirdly, while the derivation of spec-head relations in syntactic derivations is obtained by Move, which is a Copy + Merge operation, the spec-head relation falls out from M-Shift in morphological derivations.

## 3. Asymmetry in functional constructs

I now turn to the discussion of asymmetric relations in functional constructs. I bring empirical evidence to show that wh-words, complementizers and subordination conjunctions are articulated on the basis of the asymmetric relations of the M-Shell. An M-shell minimally consists of two layers of asymmetric relations, as schematized in (12), where x and y are heads, α and β are specifiers, and δ is a complement.

(12) $[_x\ \alpha\ x\ [_y\ \beta\ y\ \delta]]$

I show that +/–wh-constructs minimally include two heads in asymmetric relation with other categories. Firstly, I provide evidence for asymmetry in the upper part of the shell, i.e. the higher spec-head relation. Secondly, I motivate the projections in the lower part of the shell, the lower spec-head and head-compl relations.

## 3.1 The Op-Shell

I take wh-words to be part of operator-variable structures at LF. I propose that morpho-conceptual features they contain are structured on the basis of the Op-Shell. In the configuration in (13), the operator/variable (Op, x) and the Restrictor (y (R z)) split in two layers. The operator (Op) asymmetrically c-commands the variable (x) and the restrictor R heads a specifier-head-complement structure. The representation in (13) is independent of specific categorial features. In fact, it is part of the morpho-conceptual feature structure of all functional categories.[7]

(13) $[_{\mathrm{x}}$ Op x $[_{\mathrm{R}}$ y [R z ]]]

My proposal is compatible with the split CP approach, where different properties are represented by different projections (Ambar 2003; Cinque 1999; Rizzi 1995; Kayne & Pollock 2001; Ambar, Oberhaur et al. 1998; Munaro & Pollock 2001 among others). The point being made here is that the splitting is also part of the feature structure of wh-words and complementizers themselves. My proposal is also compatible with asymmetry based approaches to semantic and pragmatic categories, such as Stowell's (1995) analysis of Tense and Speas and Tenny's (2002) analysis of speech roles. These works further support the hypothesis that asymmetry is basic in grammar.

## 3.2 Specificity of the Op-Shell

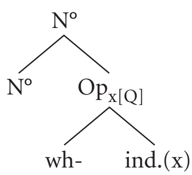
The representation in (13) differs from standard representations of the operator-variable relation, where only the operator and the variable it binds are part of the relation, as in (14). Our proposal also differs from Kamp (1981), Heim (1982), Heim and Kratzer (1994), where a quantified sentence is analyzed in terms of a tripartite logical representation, including a (raised) quantifier, a restrictive clause, and a nuclear scope, as schematized in (15a) and illustrated in (15b).

(14) $\mathrm{Op}_i$ (…$x_i$…)

(15) a. O (restrictive clause) (nuclear scope)
b. Every boy eats an ice cream.
Every$_x$ [x is a boy] ($\exists_y$) y is an ice cream $\wedge$ x eats y
Quantifier restrictive clause nuclear scope

Among the differences between (13) and (15) are the following: (13) is part of the feature structure of operators, not (15); (13) is not derived by Move as (15) is; (13) is independent of existential closure, not (15).

Our proposal is also distinct from Tsai (1994: 22), where the structure underlies wh-words, as in (16). According to his analysis, an interrogative wh-word has a Q(uestion)-operator as a binder for the variable (x) that it symmetrically c-commands.

(16)


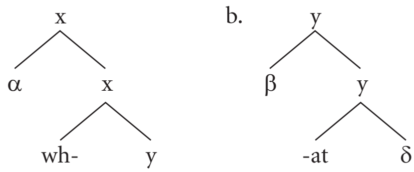

This configuration is not possible if Kayne (1994) is correct about the Linear Correspondence Axiom (LCA), which forbids adjunctions to the right. Tsai makes no commitments as to which way the Q-operator should adjoin to the wh-word. According to our proposal, the operator is originally to the left of the structure and it sister-contains (asymmetrically c-commands) the variable and the restrictor of the variable. Thus, apart from the difference in linear order, in our proposal the operator asymmetrically c-commands the variable and the complement of the variable is an articulated structure headed by the restrictor.

I now turn to the discussion of the internal articulation of the Op-Shell.

### 3.3 Two heads

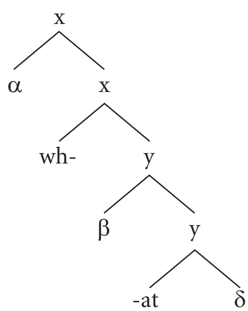
Functional constructs such as wh-words are generally formed on the basis of two morphemes in English and in Romance. In English, they include an obligatory wh- affix as well as another obligatory constituent, e.g.: *wh-o, wh-at, wh-ich, wh-ose, wh-en, wh-ere, wh-y*. Assuming that a head is obligatory in the projection it is a part of, and that each constituent of a wh-word is obligatory. I take each of the constituents of a wh-word to be a head. This is illustrated here in (17). Each head projects its features and has dependents. I will identify them below.

(17) a. b.




Given (1), the basic M-tree for wh-words is a Shell, where each head is in asymmetric relation with a spec position. The complement position must also be projected, as a spec-head relation cannot be projected without a head-compl relation. Thus, for example, the M-tree in (18) for *what* is derived by one application of M-Shift, which substitutes the root of (17b) to the complement position of (17a).[8]

(18)




Why should a wh- morpheme project a spec position? What justifies the α position in the configuration in (18)? Why is this position part of the elementary configurations in (17) to start with? I approach these questions in the perspective of Asymmetry Theory, according to which asymmetric relations are basic grammatical relations.

## 3.4 The upper layer of the shell

It is generally assumed that wh-words and complementizers act like operators that bind a variable in their local domain. I further assume, as in Chomsky (1995), that the operator feature (Q) is distinct from the wh-feature. Thus in the example in (19) from Italian, *che* is +Op in both (19a) and (19b), while it is +wh in the first case and –wh in the other.

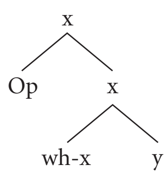
(19) a. *Che vuole*? (Italian)
What wants
'What does (she/he) want?'

b. *Credo che non lo sa.*
Think that not that know
'I think that she/he does not know.'

I take the spec position in the upper layer of the M-tree to be the locus of the operator feature, and the head of that projection to be the locus of the variable feature.

It is generally the case that heads are morphologically spelled out in M-trees, as discussed above on the basis of the properties of derivational affixes and roots. I assume that this is also the case in the internal structure of wh-words. If the wh- affix is a head, it is the morphological spell out of the variable in the M-tree. I will take the operator feature in the spec position not to be morphologically spelled out. Thus, (20) is part of the M-Shell of wh-words.

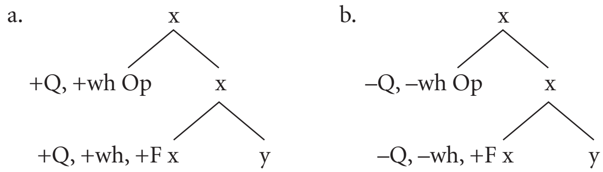
(20)



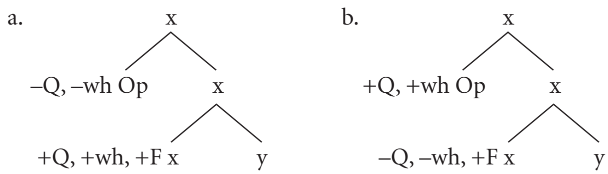

Evidence that the variable feature is asymmetrically related to the operator feature comes from independent properties of the spec-head relation. In Di Sciullo (1999: 109), I define the spec-head relation in terms of a relation between sets of features as follows:

(21) **Spec-head Feature Inclusion (agreement)**
The features of a specifier are included in the features of the head it is the specifier of.

I take the Spec-Head Feature Inclusion relation to be essentially the spec-head agreement relation. This relation holds for the operator-variable pair within the Op-Shell. It holds between the +/–Q features of the operator and the +/–wh feature of the variable. If the operator is +Q, the variable must be +wh, if the operator is –Q, the variable must be –wh. The variable is also specified for the feature F, which is not part of the features of the operator. For example, the feature +/–human that differentiates the +wh *who* from *what* in English and *chi* from *che* in Italian is an F feature.

(22)



Contrary to (22), (23) does not satisfy Feature Inclusion. The morphological spell-out of the features these structures include should not give rise to legible functional constructs.

(23)



This prediction is borne out. A specified –wh head, such as the *th-* affix in English, is a –wh variable, it cannot be part of a +Q operator Shell.[9] This prediction is borne out, as the contrasts in (24) illustrate.

(24) a. Wh-at / *Th-at is it?
b. Wh-ere / *Th-ere is it?
c. Wh-en / *Th-en is it?
d. Wh-ose / *Th-ose is it?

The hypothesis that Feature Inclusion holds for the operator-variable relation in the M-shell also makes correct predictions for Italian. If an affix is specified as –wh in that language, it will not be part of an +Q Operator M-Shell. Thus, a specified –wh head, such as *l-*, which is constant in 3rd person strong pronouns (as in *lui* 'him', *lei* 'she', *loro* 'them') and pronominal clitics (as in *lo*'him/it', *la*'her/it', *le* 'them') may not be in the sister-contain domain of a +Q spec. The examples in (25) illustrate this point .

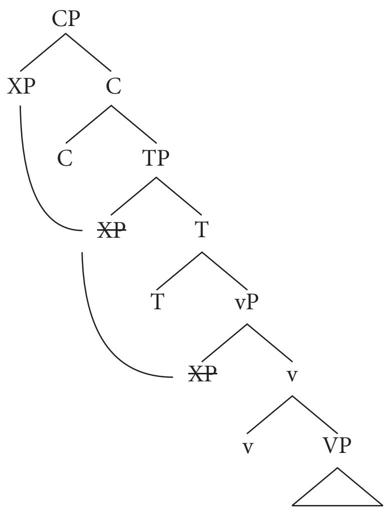
(25) *Ch-e /*L-e fare*? (Italian)
What / that do
'What to do?'

(26) *Qu-oi /*L-oi faire*? (French)
What / that do
'What to do?'

Thus, in an M-tree the features of the head are in asymmetric relation with the features in spec, and Feature Inclusion holds for the operator-variable pair. This allows for correct predictions to be made with respect to the restrictions on the derivation of functional constructs.

M-trees for wh-words are not isomorphic to S-trees with wh-words. In Chomsky (1995) the operator/variable relation is obtained derivationally by overt or covert movement. Wh- expressions raise to the specifier of a functional projection (CP or higher projections) either overtly or covertly. The syntactic operator/variable relation is thus the result of movement, as schematized in (27).

(27)



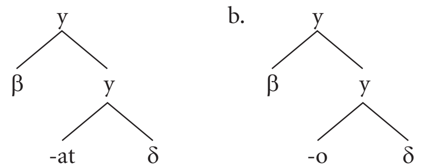

Assuming the copy theory of movement, according to which movement is Copy + Merge, the variable bound by an operator is no longer available in terms of a trace. It can however be derived by lambda abstraction, along the lines of Fox (2002), in which case the derived representation is still formally distinct from the M-tree. In the M-tree, the variable in the head position is adjacent to the operator in the spec position, whereas the variable is not adjacent to the operator in the S-tree. A further difference lies in the PF visibility of the operator and the variable. In an M-tree, the head is generally PF visible and the operator feature in spec is not, whereas in an S-tree, the variable is generally not PF visible, whereas the operator feature generally is. Finally, contrary to syntax, the Q feature must not be deleted in M-derivations, as the value of Q is determinant in Feature Inclusion in M-trees.[10]

In the next section, I consider the asymmetry in the lower part of the M-shell. It will become clear that the spec-head asymmetry also holds for the constituents of the lower layer, and that part of this asymmetry is not visible to the syntactic plane.

## 3.5 The lower layer of the shell

The lower layer of the M-shell is isomorphic to its upper layer. It is constituted of a head asymmetrically related to a spec position. In the case of wh-words, the head of the lower layer of the Shell is not a constant morpheme. In most cases it is a syllable, as with the *-at* of *what*, and with the *-o* of *who*. According to our proposal, these elements head structures such as (28a) and (28b).

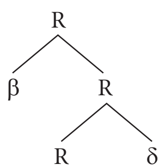
(28) a. b.




How could such impoverished morphemes head a projection? Why should they project a spec position? What justifies this position? If the compl position of the upper layer of the M-shell is justified because it is the target of the operation M-Shift, what justifies the lower compl position?

The justification comes from a property of the head of the lower layer of the Op-Shell, relating δ to β. This is schematized in (29), where the head of the projection is a relational element R.

(29)




I start by qualifying the relational property of the head in the lower layer of the Shell, which, according to my hypothesis, compositionally restricts the variable in the upper layer. In my view, the restrictor of the variable in the M-plane is not a feature, nor a list of features, but rather it is a structured set of features.

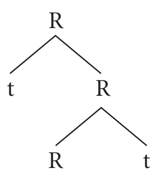
It has been observed that the morphological spell-out of wh-words differs according to the features of the restrictor of the variable. The restrictor ranges over categorial features, Case and argument/adjunct features. Thus, Tsai (1994), following Huang (1982), distinguishes nominal wh-phrases from non-nominal ones with respect to extraction. He argues that *where* and *when* are nouns, as opposed to the adverbs *how* and *why*. *Where* and *when* pattern like arguments with respect to LF locality effects. Thus, some wh-words can only be interpreted as arguments, as it is the case for *who* and *what*, while others can only be interpreted as adjuncts, as it is the case for *how* and *why*. It has also been assumed that semantic features such as the +/–human feature are part of the restrictor of the variable and differentiate wh-expressions such as *who* and *what*. I referred to these features as F features above, and assumed that they were part of the features of the variable in the upper layer of the Op-Shell.

There is a basic property of wh-words that is not captured under an analysis where the restrictor of the variable is reduced to a feature or a list of features. This property is that a wh-word is a relational element. Thus, a +wh element turns an unsaturated proposition onto a set of propositions.[11] This is illustrated in (30).

(30) a. Q: What happened?
b. $A_i$: John came in. $A_j$: Paul went out. $A_k$: Mary stayed in,. . ., $A_n$
c. Q: Who did Mary see?
d. $A_i$: Mary saw John. $A_j$: Mary saw Lucy and Debby. $A_k$: Mary saw Juliet,..., $A_n$
e. Q: When did John arrive?
f. $A_i$: John arrived at noon. $A_j$: John arrived late. $A_k$: John arrived in the morning,. . ., $A_n$

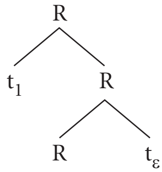
I propose that the relational property of functional constructs is part of the lower layer of the M-shell. In the case of +wh constructs, the dependents of R are propositional types of variables (t), as in (31).

(31)




Asymmetry holds in the lower layer of the Op-Shell. In the case of +wh words, the feature in spec is the feature of an unsaturated proposition $t_1$. The feature in compl is the feature of the set of possible answers $t_\varepsilon$ to the question. Thus, the following M-tree for the lower layer of +wh words is obtained.

(32)




The examples in (33) and (34) show that a –wh word turn a proposition onto another proposition.

(33) a. I know *that* snow is white.
b. I thought *that* snow was white.
c. I expect *for* John *to* be on time.
b. I wonder *if/whether* John will come.

(34) a. John left *as if* he was bored by the meeting.
b. Paul will arrive today *if not* he will arrive tomorrow.
c. Luc came in *then* he climbed the stairs.

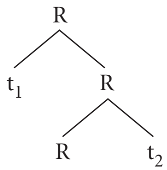
–Wh elements can be thought of as coordination operators that relate two entities of the same type.[12] The morphological spell-out of –wh expressions differs whether the restrictor relates two propositions, the truth value of which could be true, as in (33a), false, as in (33b), not determined, as in (33c), or either true or false, as in (33d). Evidence that –wh elements relate truth conditional elements comes from the fact that in Italian as well as in the other Romance languages including French, the conjunction itself may include the spell-out of the truth value 'true' and 'false' in the form of the adverbials 'yes' and 'no'. See the examples in (35) from Italian and (36) from French.

(35) a. *La neve è bianca siccome la neve è bianca.* (Italian)
The snow is white yes as the snow is white
'Snow is white as snow is white.'
b. *La neve è bianca sennò siamo nel mondo B.*
The snow is white if not are in the world B
'Snow is white otherwise we are in world B.'

(36) a. *Il ne sait si oui ou non il partira.* (French)
He not know if yes or no he will leave
'He does not know whether he will leave.'
b. *Il partira si non je partirai.*
He will leave if not I will go
'He will leave if not I will.'

These facts bring empirical evidence to show that the –wh operator Shell includes a restrictor of the variable, and that the restrictor relates truth conditional elements. This is schematized in (37).

(37)




I thus take a –wh element to be a conjunction operator, the head of which projects the typical asymmetric structure of the M-shell. The relational property of the conjunction operator is expressed in terms of the relation R, the head of the restrictor layer. Thus, the restrictor layer of +wh and –wh words differs minimally in feature structure, as can be seen by comparing (32) and (37). A +wh restrictor relates a proposition to a set of propositions, whereas a –wh restrictor relates a proposition to another proposition.

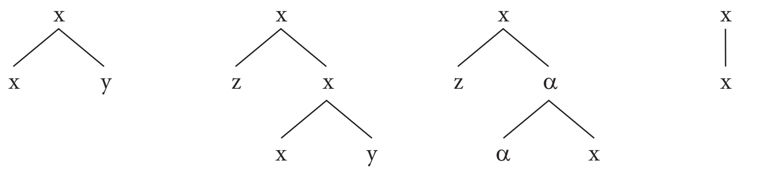
Furthermore, a spatio-temporal precedence relation holds between the constituents of the lower layer of the Op-Shell.[13] With +wh words $t_1$ precedes $t_\varepsilon$ , and with –wh words $t_1$ precedes $t_2$. Empirical evidence that this is the case comes from the fact that the inverse order of the propositions affects legibility. In the case of +wh, it gives rise to echo questions, see (39), and in the case of –wh it gives rise to gibberish, see (40).

(39) a. A: John arrived at noon.
Q: When did John arrived?
b. A: Mary saw John.
Q: Who did Mary see?

(40) a. #Snow is white I know.
b. #John to be on time I wonder.
c. #Mary will come I don't know.

+/–Wh elements express relations between propositions, that is, relations between sets of truth values. They also express relations between events, that is, relations between ordered sequences of propositions in time and in abstract space. I identified this relation as the Path relation in Di Sciullo (2000).[14,15] Prepositions are the core elements in the Path relation. They have been classified with respect to their denotation of spatial relations. They also have been analyzed in Hale (1998: 2) as the canonical relational categories, as depicted here in (41) for English. They are most likely candidates to be part of the Op-Shell. In fact they are as evidenced in (42) with prepositional complementizers.

(41)




x = V x = P/V x = A x = N

(42) a. *Gianni ha cominciato <u>a</u> lavorare.*
'Gianni started to work.'
b. *Gianni ha smesso <u>di</u> lavorare.*
'Gianni stopped working.'

This brings empirical support to our analysis according to which the lower layer of the Shell includes a relational head, which may be thought of as being encored in the event structure in terms of the asymmetric Path relation.

## 4. The full Shell

In the previous sections, I proposed that the semantic features in +/–wh words are structured on the basis of the asymmetric relations of the M-Shell. In the configuration in (43), the operator/variable relation (Op, x) and the Restrictor ((y, R) (R, z)) relations split in two layers. The operator (Op) sister-contains the variable (x) and the restrictor R heads a specifier-head-complement structure, where it compositionally restricts the variable in the domain of the operator. The representation in (43) is independent of specific categorial features. In fact, it is part of the feature structure of functional categories, including determiners and demonstratives. My proposal provides a unified analysis of the properties of +/– wh words.

(43) $[_{xOp}\ \mathrm{Op}\ x\ [_{R}\ y\ [R\ z\ ]]]$

(44) $[_{+wh}\ \mathrm{Op}_{+Q}\ x\ _{+wh}[_{R}\ t_1[\ R\ t_{\varepsilon}]]]$

(45) $[_{-wh}\ \mathrm{Op}_{-Q}\ x_{-wh}\ [_{R}\ t_1[\ R\ t_2]]]$

The representations in (44) and (45) differ with respect to the terms in the specifier and the complement of R. In the case of the +wh feature, R relates an unsaturated proposition $t_1$ in the specifier of R to a set of propositions $t_{\varepsilon}$ in the complement of R, which is the set of possible answers to the wh-question. In the case of the –wh feature, R relates an unsaturated proposition $t_1$ in the specifier of R to another proposition $t_2$, in the complement of R. As it is the case for the upper part of the Shell, the lower part of +wh and –wh Shells differs only in feature specification. The upper part of the Shell also differs minimally in the specific features of the operator and the variable. Notwithstanding the featural differences, the M-shells deploy the asymmetry of morphological relations. In particular, every head is in asymmetric relation with another constituent.

The morphological asymmetries cannot be equated with the syntactic asymmetries. Thus, the operator-variable asymmetry in the syntactic plane is an A-bar/A (spec-spec or spec-compl) asymmetry, it is not a spec-head asymmetry, as it is the case in the morphological plane.

## 5. Adjuncts

The discussion above brings evidence to the effect that asymmetry holds in both layers of the M-Shell. In this section, I show that the asymmetry of morphological relations extends outside the minimal domain of the M-shell. To do so, I consider the properties of +wh adjuncts as well as adverbial subordination conjunctions. I will start by considering the properties of +wh adjuncts, and the structure of *why* in particular.

In Romance, a preposition is part of the equivalent +wh word for *why*, e.g.: *perché* (Italian), *pourquoi* (French), *porqué* (Spanish), *porque* (Portuguese), *pentru ce* (Romanian). This preposition is covert in the English form, but present nevertheless, as it is part of the possible answers to *why* questions, see (46).

(46) a. Why did they leave so early?
  b. $A_i$: To be on time for their flight.; $A_j$: For John to be on time.; $A_k$ For no reason.,. . ..., $A_n$

How is the preposition related to the rest of the configuration? The following facts from Romance are telling in this respect. Firstly, the preposition is not an obligatory element, as the expressions are still +wh words without the presence of the preposition, see (47). Secondly, the preposition is constant, while the remaining part of the expression may not be, see (48). Thirdly, the presence of the preposition in the +wh words does not affect the argument structure of the proposition it is a part of, whereas the absence of the preposition does, see (49).

(47) a. *Perchè vuoi vedere Gianni?*
Why want to see Gianni
'Why do you want to see Gianni?

b. *Che voi vedere?*
What want to see
'What do you want to see?'

(48) a. *perchè*, *perchi* (Italian)
'why', 'for who'

b. *pourquoi*, *pour qui* (French)
'why', 'for who'

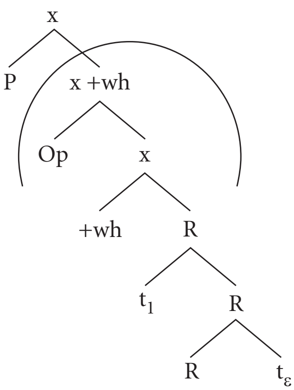
c. *pentru ce*, *pentru cine* (Romanian)
'why', 'for who'

(49) a. *Perchè Gianni ha colpito Mario*?
For what Gianni hit Mario
'Why did Gianni hit Mario?'

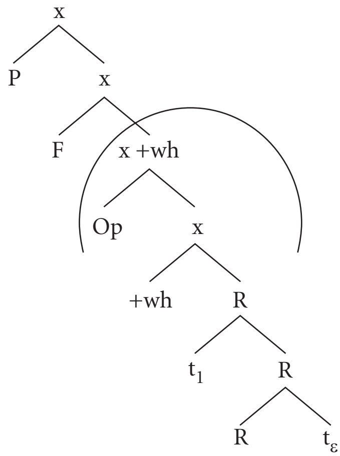
b. **Chè Gianni ha colpito Mario*?
What Gianni hit Mario
'What did Gianni hit Mario?'

Given the facts above, I analyze the preposition *per* in *perché* as an adjunct to the Op-Shell of *che*. A possible implementation of this analysis would be to Chomsky-adjoin the preposition to the root of the Op-Shell, as in (50). Another possibility would be to take the adjunct to be in spec of a functional head F, along the lines of Cinque (1999), as in (51). I will not provide arguments to prefer one analysis over the other here, but see Di Sciullo (forthcoming) for discussion. In the representation in (51), the F head is in asymmetric relation with the preposition in spec, thus satisfying Head-Asymmetry. In the representation in (50), the prepositional adjunct sister-contains the operator, as well as all the other nodes in the domain of the operator.

(50)

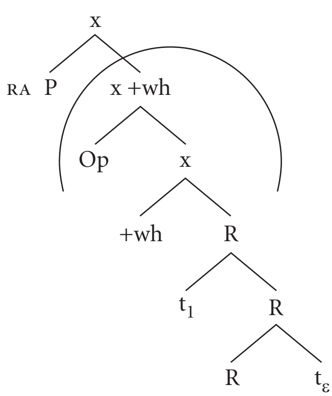



(51)

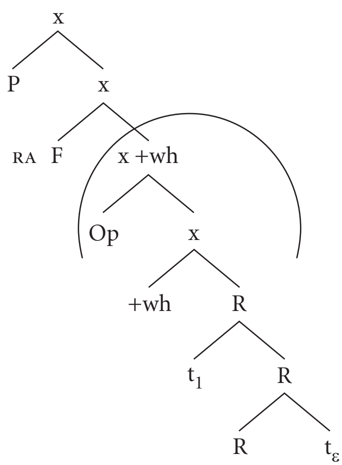



Here again asymmetry holds. It holds from the features of the prepositional adjunct to the features of the Op-Shell, locally in (50) and via the F head in (51). In the case at hand, the relevant feature of the adjunct is the RATIONALE feature (RA). This feature is outside the domain of the minimal Shell, as depicted in (52) and (53).

(52)




(53)




The adjunct RA feature identifies by Linking the propositional feature $t_\varepsilon$ of the lower layer of the shell in the case of a +wh construct, and it identifies by

Linking the propositional feature $t_2$ in the case of a –wh construct. The Linking targets, via the +wh head, the R head of the lower layer of the Op-Shell.

This correctly predicts that the RA feature may not modify non propositional operators, such as pronouns, e.g. *lui* 'him', *lei* 'her', *loro* 'them', which may otherwise be in the domain of a preposition. This prediction is borne, see (54) and (55).

(54) a. *Per-ché/*per-lui hai fatto questo?*
For what/ For him have done this?
'Why did you do this?'

b. *Per-chi /*per-lei hai fatto questo*?
For who/For her have done this?
'For who did you do this?

(55) *L'ho fatto per-ché/*per-loro dovevo farlo.*
It have done for that/ for them must do it
'I did it because I had to.'

Further evidence comes from the fact that in Italian, and more generally in Romance, adverbial prepositions are part of subordination conjunctions, such as *primaché* 'before', *dopoché* 'then', *poiché* 'as', *allorché* 'when', where an adverbial constituent is part of the –wh word, as illustrated in (56).

(56) a. *Gianni è uscito primaché Paulo sia entrato.*
'Gianni went out before Paulo came in.'

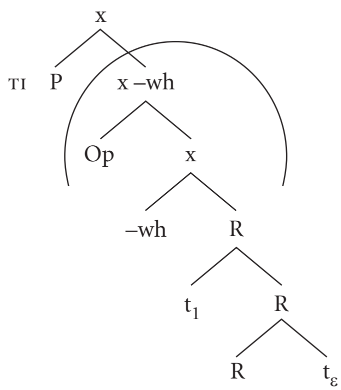
b. *Gianni è uscito dopoché Paulo è entrato*.
'Gianni went out afterwards Paulo came in.'

c. *Gianni è uscito poiché Paulo è entrato*.
'Gianni went out as Paulo came in.'

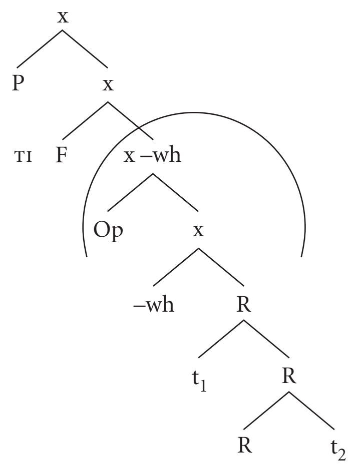
d. *Gianni è uscito allorché Paulo è entrato*.
'Gianni went out while Paulo came in.'

These –wh expressions are conjunction operators. They relate two propositions and are structured on the basis of the operator Shell. The Shell includes a conjunction-operator (Op) sister-containing a variable (x) and a restrictor structure where the specifier and the complement of R are propositions ($t_1$ and $t_2$). The temporal adverb (TI) is an adjunct to the M-Shell, as depicted below.

(57)




(58)




The –wh connective expresses a temporal relation between two propositions, where $t_2$ occurs in a subsequent, precedent or identical point in time than $t_1$.

Here again asymmetry holds. It holds from the features of the adverbial adjunct to the features of the Op-Shell. In the case at hand, the relevant feature of the adjunct is the TI feature. This feature is external to the minimal Shell. The TI adjunct identifies the R head via the –wh head sister-contained by a non propositional Op such as a demonstrative, e.g. *questo* 'this-masc.', *questa* 'this-fem.', *questi* 'these'. This correctly predicts that the temporal adjunct may not modify a non-propositional operator, as can be seen by comparing the examples in (56) above with the examples in (59).

(59) a. **Gianni è uscito prima-questo Paulo sia entrato.*
'Gianni went out before this Paulo came in.'
b. **Gianni è uscito dopo-questa Paulo è entrato.*
'Gianni went out afterwards this Paulo came in.'
c. **Gianni è uscito poi-questi loro è entrato.*
'Gianni went out as these they came in.'
d. **Gianni è uscito allora-questa lei è entrato*
'Gianni went out while-this she came in.'

Temporal and rationale adjuncts identify the conceptual pseudo-space according to which the +/–wh receives an interpretation. The asymmetry of adjuncts with respect to the M-shell is expected in Asymmetry Theory and is not identical to the syntactic asymmetry of adjunct wh-phrases.

## 6. Summary

In this paper, I provided empirical justification for the hypothesis that +/–wh features are articulated on the basis of the asymmetry of the M-Shell, where Head-Asymmetry holds for the operator-variable relation as well as for the variable-restrictor relation. As expected, the properties of functional Shells are not isomorphic to the properties of functional projections in syntactic structures.

Contrary to common assumptions, it appears that M-Shells are more restricted than the syntactic projection of functional heads. In effect, the locality conditions, as well as the PF visibility conditions are more severe in morphological objects than in syntactic objects. Moreover, differences are observed in the linear order of constituents and with respect to visibility which also suggests that morphological and syntactic objects, even though related, cannot be equated.

My proposal has consequences for the properties of the interfaces between morphological structure and the Conceptual-Intentional system. It brings further evidence in support of the fact that the grammar ensures distinctiveness between morphological and syntactic relations at a minimal cost. This cost is in effect minimal if asymmetry is a fundamental property of the relations shared by the different sorts of objects derived by the grammar.

## Notes

**1.** The material presented here is further elaborated in the forthcoming MIT Press monography *Asymmetry in Morphology*. This work is supported in part by the Social Sciences and Humanities Research Council of Canada to the Major Collaborative Research Project on Asymmetries in Natural Languages and their Treatment by the Performance Systems (Grant no. 412-97-0016). www.asymmetryproject.uqam.ca. Many thanks to the members of the asymmetry project. Special thanks to Manuela Ambar, Greg Carlson, Noam Chomsky, Manuel Espanol Echevarria, Jacqueline Guéron, Ken Hale, Jim Higginbotham, Jay Keyser, Jean-Yves Pollock, James Pustejovsky, Dana Isac, Richie Kayne, Andrea Moro, Tom Roeper, Stanca Somesfalean, Peggy Speas, Ed Stabler, Carol Tenny, Eric Wehrli and Edwin Williams for fruitful discussions.

**2.** See also Hale and Keyser (1993) and Keyser and Roeper (1997), on the properties of lexical relations; Keyser and Roeper (1997), Halle and Marantz (1993), Roeper (1999) for the properties of morphological relations; Epstein (1995), Frank and Vijayashankar (1995) and Reuland (1998) on the properties of syntactic relations, Piggott and Hulst (1997), Dresher and Hulst (1998), Hulst and Ritter (1999), Raimy (2000), for the properties of phonological relations.

**3.** This supports the hypothesis that words and phrases are different types of grammatical objects. Morphological objects differ from syntactic objects formally and semantically. Contrary to syntactic objects, morphological objects have a unique stress, their constituents are not separable and they are not fully compositional. Moreover, these objects exhibit asymmetry differently. See Di Sciullo and Williams (1987), Williams (1994), Bach (1996), Di Sciullo and Tenny (1997), Di Sciullo (2000) among others.

**4.** An asymmetric relation holds for the ordered pairs in a set if the pair < x, y> is part of that set and not the pair < y, x>. Considering structural relations, asymmetric c-command or the sister-contain relation, in the sense of Chomsky (1998) is asymmetric. In Chomsky (1998) c-command falls out from the computational process. Thus, the operation Merge takes two elements α and β and forms a more complex one K constructed from α, β. Merge provides two relations: sisterhood, which holds for (α, β) and immediate contain , which holds for (K, α) and (K, β). By composition of relations two new relations are derived: the relation contain and the relation c-command (sister-contain). Thus, K contains α if K immediately contains α, or K immediately contains L which immediately contains α; conversely α is a term of K if K contains α. And α c-commands β if α, is the sister of K which contains β. See Chomsky (1998) for discussion.

**5.** An antisymmetric relation is asymmetric and reflexive, whereas an asymmetric relation is irreflexive. Contrary to an asymmetric relation, an antisymmetric relation may include reflexive relations such as <x, x> and <y, y>.

**6.** Symmetry is a property of the relations in a set such that if the pair <x, y> is part of that set, the pair <y, x> is also. The 'sister of' relation qualifies as symmetric, as if x is the sister of y, y is also the sister of x.

**7.** My proposal extends to the morpho-conceptual properties of the other functionals including D elements, such as determiners, demonstratives, pronouns and clitics. See Di Sciullo (2000, forthcoming) for discussion.

**8.** I assume that the elementary trees in (17a) and (17b) are part of the lexicon. Functional affixes do not differ from derivational affixes in this respect. See Di Sciullo (1999b) for discussion.

**9.** Here are some Op-Shells for –wh D Operators. In these representations, the –wh variable in the upper layer is spelled out by *th-*, the relational head in the lower layer is specified for spatial features, +/–Proxi(mate), and +Loc(ation). The head of the lower clause relates a predicate variable (pred) to an individual type of variable e. See Di Sciullo (2000, forthcoming) for discussion.

**10.** According to Chomsky (1995–1998), Lasnik (1999) and related works, the wh- feature is an interpretable functional feature. Thus, this feature must be visible at LF. This is not the case for uninterpretable functional features, such as the Q feature. Uninterpretable features must be checked/erased before Spell-Out, and thus are not visible for LF interpretation.

**11.** In the various theories of the semantics of questions, as is the case in Kartunnen (1977), all wh-expressions including wh-morphemes such as *who, what,* and *which* are uniformly rendered as existential quantification over sets of possible answers ($A_i$, $A_j$, $A_k$,...,$A_n$).

**12.** The asymmetrical nature of the structural relations underlying coordination phrases is also argued for in Munn (1992, 1993), Kayne (1994), and Johannessen (1996, 1998), where a conjunction such as *and* is a functional element that heads an asymmetric specifier-head-complement projection. See Di Sciullo (2000).

**13.** The properties of Event structure have been discussed widely in the literature. See Bach (1996); Binnick (1998); Carlson (1977, 1999); Comrie (1976); Hale (1984); Kipka (1990); Klipple (1991); Higginbotham (1999a,b); O'Keefe (1996); Parson (1990); Talmy (1985, 1996); Tenny (1987, 1994); Vendler (1967); Verkuyl (1972, 1993), among others.

**14.** In Vector Grammar (O'Keefe 1996: 279), spatial relations are specified in terms of places, directions and distances. A path is an ordered sequence of places and a 'translation vector' between them. Translations of positions in a space are denoted by vectors whose tail begins at the origin of movement and whose head ends at the destination.

**15.** I proposed that R has a spatio-temporal encoring and that this encoring can be expressed in terms of a Path, that is a relation between two points in a space. I used the following representation [ Terminus [ P Initial Point]] to express the path relation within +/–wh- Op-Shell, the Terminus is $t_\varepsilon$ for +wh and $t_2$ for –wh, and the Initial Point is $t_1$ in both cases. I will not discuss the details of this analysis here. See Di Sciullo (2000, forthcoming).

## References


Ambar, M. (2003). Wh-asymmetries. In A. M. Di Sciullo (Ed.), *Asymmetry in Grammar*, Vol. 1: *Syntax and Semantics*. Amsterdam/Philadelphia: John Benjamins,

Ambar, M., Obenhauer, H. et al. (1998). From Wh-questions to Wh-interrogatives: the internal structure of Wh-phrases and the left periphery. Evidence from Portuguese, French and Hungarian. Article posted at Chomsky Celebration. *Internet*.

Bach, E. (1996). On the grammar of complex words. In *Configurations. Essays on Structure and Interpretation*. A. M. Di Sciullo (Ed.), Somerville, MA: Cascadilla Press.

Borer, H. (1991). The causative-inchoative alternation. A case study in parallel morphology. *The Linguistic Review, 8*.

Binnick, R. (1998). *Time and the Verb*. Oxford University Press.

Carlson, G. (1977). Reference to kinds in English. Doctoral dissertation. University of Massachusetts, Amherst.

Carlson, G. (1999). Event semantics and propositional semantics. Termi Conference on Tense, Aspect and Mood. Termi, Greece.

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (1998). *Minimalist Inquiries*. Ms., MIT.

Chomsky, N. (1999). *Derivation by Phase*. Ms., MIT.

Collins, C. (1997). *Local Economy*. Cambridge, MA: MIT Press.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Di Sciullo, A. M. (1996). X-bar selection. In J. Rooryck and L. Zaring (Eds.), *Phrase Structure and the Lexicon* (pp. 77–107). Dordrecht, Kluwer.

Di Sciullo, A. M. (1999a). Formal context and morphological analysis. In Paulo Bouquet, Luciano Serafini, Parick Brézillon, Massimo Benerecetti, Francesca Castellani (Eds.), *Modelling and Using Context* (pp. 105–119). Springler.

Di Sciullo, A. M. (1999b). The local asymmetry connection. In *Papers from the UPenn/MIT Roundtable on the Lexicon. MIT Working Papers in Linguistics, 35*, 25–49.

Di Sciullo, A. M. (2000). Paths in words. In *ESSLI* 2000. Birmington.

Di Sciullo, A. M. (Forthcoming). *Asymmetry in Morphology*. Cambridge, MA: MIT Press.

Di Sciullo, A. M. & Tenny, C. (1997). Modification, event structure and the word/phrase asymmetry. *NELS* 28. GLSA, University of Massachusetts, Amherst.

Di Sciullo, A. M. & Williams, E. S. (1987). *On the Definition of Word*. Cambridge, MA: MIT Press.

Dresher, E. & van der Hulst, H. (1998). Head dependency in phonology: complexity and visibility. *Phonology, 15*, 317–352.

Embick, D. (2000). Features, syntax, and categories in the Latin perfect. *Linguistic Inquiry, 31*, 185–231.

Epstein, S. (1995). The derivation of syntactic relations. Ms., Harvard University.

Fox, D. (2002). Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry, 33*, 63–97.

Frank, R. & Vijayashankar, K. (1995). C-command and grammatical primitives. *Glow 1995 Newsletter.*

Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA: MIT Press

Hale, K. (1984). Notes on world view and semantic categories: some Walpiri examples. Ms., MIT.

Hale, K. (1998). Universal grammar and the necessity of linguistic diversity. Ms., MIT.

Hale, K. & Keyser, J. S. (1993). On argument structure and the lexical expression of syntactic relations. In Kenneth Hale and Jay S. Keyser (Eds.), *The View from Building, 20* (pp. 53–110). Cambridge, MA: MIT Press.

Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In Ken Hale and Jay Keyser (Eds.), *The View from Building 20: Essays in Linguistics in Honnor of Sylvain Bromberger* (pp. 111–176). Cambridge, MA: MIT Press,

Heim, I. (1982). The semantics of definite and indefinite noun phrases. Doctoral dissertation, University of Massachusetts.

Heim, I. & Kratzer, A. (1994). *Semantics in Generative Grammar*. Malden, Mass.: Blackwell.

Higginbotham, J. (1999a). On events in linguistics semantics. In J. Higginbotham, J. Pianesi, and A. Varzi (Eds.), *Speaking of Events* (pp. 145–176). Oxford University Press.

Higginbotham, J. (1999b). On the representation of telicity. Termi Conference on Tense, Aspect and Mood. Termi, Greece.

Huang, J. (1982). Logical relations in Chinese and the theory of grammar. Doctoral dissertation. MIT.

Hulst, H. & Ritter, N. A. (1999). Head driven phonology. In H. van der Hulst and N. A. Ritter (Eds.), *The Syllabe: Views and Facts* (pp. 113–168). Berlin Moutin de Gruyter.

Isac, D. (2003). Restrictive relative clauses. In A. M. Di Sciullo (Ed.), *Asymmetry in Grammar*, Vol. 1: *Syntax and Semantics*. Amsterdam/Philadelphia: John Benjamins.

Johannessen, J. B. (1996). Partial agreement and coordination. *Linguistic Inquiry, 27* (4), 661–76.

Johannessen, J. B. (1998). *Coordination.* Oxford: Oxford University Press.

Kamp, H. (1981). A theory of truth and semantic representation., In J. Groenendijk, T. Janssen and M. Stokhof (Eds.), *Formal Methods in the Study of Language* (pp. 277–321). Amsterdam: Mathematical Center.

Kartunnen, L. (1977). Syntax and semantics of questions. *Linguistic and Philosophy, 1*, 3–44.

Kayne, R. (1994). *The Antisymmetry of Syntax.* Cambridge, MA: MIT press.

Kayne, R. & Pollock J.-Y. (2001). New thoughts on stylistic inversion. In A. Hulk & J.-Y. Pollock (Eds.), *Subject Inversion and the Universal Grammar*. New York: Oxford University Press.

Keyser, J. & Roeper, T. (1997). Anti-symmetry and leftward movement in morphology. Ms., MIT and University of Massachusetts.

Kipka, P. (1990). Slavic aspect and its implications. Doctoral dissertation, Cambridge, MA: MIT.

Klipple, E. (1991). The aspectual nature of thematic relations. Doctoral dissertation, MIT.

Larson, R. (1987). Missing prepositions and English free relative clauses. *Linguistic Inquiry, 18*, 239–267.

Law, P. (1997). Modularity in nominalizations. In A. M. Di Sciullo (Ed.), *Conditions and Interface Projections* (pp. 28–51). Oxford: Oxford University Press.

Lieber, R. (1980). On the organization of the lexicon. Doctoral dissertation, MIT.

Moro, A. (2001). *Dynamic Antisymmety*. Cambridge, MA: MIT Press.

Munn, A. B. (1992). A null operator analysis of ATB gaps. *The Linguistic Review, 9*, 1–26.

Munn, A. (1993). Topics in the syntax and semantics of coordinate structures. Doctoral dissertation, University of Maryland.

Murano, N. & Pollock, J.-Y. (2001). Qu'est-ce que 'qu'est-ce que', a case study in comparative Romance interrogative syntax. Unpublished article, Venice & Amiens: to appear in Kayne & Cinque (Eds.), *Comparative Syntax*. Oxford University Press.

O'Keefe, J. (1996). The spatial prepositions in English, vector grammar, and the cognitive map theory. In *Language and Space* (pp. 277–316). Cambridge, MA: MIT Press.

Parson, T. (1990). *Events in the Semantics of English: A Study of Subatomic Semantics*. Cambridge, MA: MIT Press.

Partee, B., Ter Meulen, A. & Wall, R. (1990). *Mathematical Methods in Linguistics*. Dordrecht, Kluwer Academic Publisher.

Piggott, G. L. & H. van der Hulst (1997). Locality and the nature of nasal harmony. *Lingua, 103*, 85–112.

Raimy, E. (2000). Remarks on backcopying. *Linguistic Inquiry, 31*, 541–552.

Reuland, E. (1998). Deriving C-command in binding. *NELS, 18*.

Rizzi, L. (1990). *Relativized Modularity*. Cambridge. MA: MIT Press.

Roeper, T. (1999). Leftward movement in morphology. Ms., University of Massachusetts.

Roeper, T. & D. Siegel (1978). A lexical transformation for verbal compounds. *Linguistic Inquiry, 9*, 197–260.

Selkirk, E. (1982). *The Syntax of Words*. Cambridge, MA: MIT Press.

Speas, P. & Tenny, C. (2003). Configurational properties of point of view roles. In A. M. Di Sciullo (Ed.), *Asymmetry in Grammar*, Vol. 1: *Syntax and Semantics*. Amsterdam/Philadelphia: John Benjamins.

Stowell, T. (1995). The phrase structure of tense. In J. Rooryck and L. Zaring (Eds.), *Phrase Structure and the Lexicon*. Dordrecht, Kluwer.

Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (Ed.), *Language Typology and Syntactic Description*. Vol. 3., Grammatical Categories and the Lexicon, 47–149. Cambridge: Cambridge University Press.

Talmy, L. (1996). Fictive motion in language and 'ception. In *Language and Space* (pp. 211–276). Cambridge, MA: MIT Press.

Tenny, C. (1987). Grammaticalizing aspect and affectedness. Doctoral dissertation, MIT.

Tenny, Carol. (1994). *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer Academic Publishers.

Tsai, W.-T. D. (1994). On economy and the theory of A-bar dependencies. *MITWPL*.

Vendler, Z. (1967). *Linguistics in Philosophy*. Ithaca, N.Y.: Cornell University Press.

Verkuyl, H. (1993). *A Theory of Aspectuality*. Cambridge: Cambridge University Press.

Wall, R. (1972). *Introduction in Mathematical Linguistics*. Englewood Cliffs, New Jersey: Prentice Hall.

Williams, E. (1981). On the motions 'lexically related' and 'head of a word.' *Linguistic Inquiry, 12*, 245–274.

# Prefixation vs. compounding

## The case of Greek preverbs*

Angela Ralli
University of Patras

## 1. Introduction

In this paper, I deal with the role of three Greek preverbs, *ksana, kse-*, and *para-* in verb formation. By examining their characteristics on semantic, structural, and phonological grounds, I propose a morphological approach that allows for a unified treatment of the three preverbs and accounts for both their differences and similarities. I show that a distinction into internal and external preverbs, or a division in prefixes and non-prefixes, are not sufficient to take into consideration their peculiar behavior. I argue that the combination of the preverbs with the root verbs is an instance of a word-formation process that is handled within a morphological module of grammar and refers to either derivational prefixation or compounding. Following Di Sciullo 1997, I consider that the combination of the preverbs with the verbal bases occurs at different adjunction sites in grammar. However, I propose that these adjunction sites can be expressed in terms of morphological categories, i.e., as adjunction to stems or adjunction to words. I show that a difference in adjunction sites within morphology can explain the peculiar character of certain preverbs that behave like prefixes but share more similarities with word-like preverbs, than with prefix-type ones.

## 2. General characteristics

Preverbs constitute a major issue in the grammar of the Greek language. They pose an interesting puzzle with respect to the following properties:[1]

a. There are preverbs that are structurally bound morphemes, i.e., prefixes, while others behave like free elements, e.g., they share features with adverbs.
b. The preverb-verb formations display properties of phonological words, independently of the specific character of each preverb to behave like a bound or a free morpheme.
c. Preverbs select a verb to combine with, but do not determine the category of the projections they are part of.
d. Some preverbs bring only external specifications to the root meaning of the verb, while there are also preverbs affecting its aktionsart, its argument structure, or may cause a complete change of the root meaning.

*Ksana*, *kse*- and *para*- are frequently used in Modern Greek, but are of distinct historical development and different behavior. *Ksana* and *kse*- appear for the first time in late Medieval Greek with more or less the same characteristics as today (see Gardikas 1924; Hatzidakis 1926; Karantzola & Giannoulopoulou 2000).[2] *Para*- has an Ancient Greek origin, but as shown in (1), has developed some additional features that can also be traced back in late Medieval Greek. That is why *ksana* and *kse*- are usually adjoined to common verbs of Modern Greek, i.e., to verbs that are not marked as [learned], while *para*- appears with verbs of both types, that is with verbs that are marked and verbs that are not marked as [learned]:[3]

| (1) | Verb | *para*-[4] | *ksana* | *kse*- |
|---|---|---|---|---|
| | a. *fuskono* | *parafuskono* | *ksanafuskono* | *ksefuskono* |
| | to inflate | to over-inflate | to re-inflate | to deflate |
| | b. *kampto* [*learned*] | *parakampto* | **ksanakampto* | **ksekampto* |
| | to bend | to deviate | to rebend | to unbend |

Furthermore, verbs presenting a form alternation, i.e., a modern form (2b,d) and a form that is (2c) or looks like ancient (2a), freely accept *para*-, while *ksana* and, *kse*- are adjoined only to the new form which is not marked as [learned]:

| (2) | Verb | *para*- | *ksana* | *kse*- |
|---|---|---|---|---|
| | a. *δiδo* | *paraδiδo* | **ksanaδiδo* | **kseδiδo*[5] |
| | b. *δino* | *paraδino* | *ksanaδino* | *kseδino* |
| | to give | to deliver | to re-give | to relax |
| | c. *lamvano* | *paralamvano* | **ksanalamvano* | |
| | d. *laveno* | *paralaveno* | *ksanalaveno* | |
| | to take | to receive | to retake | |

The three preverbs do not behave uniformly as sentential elements. *Ksana* is the adverb for expressing repetition in Modern Greek, and is present either as an adverbial word-form in sentences (3a), or as a preverb in verbal formations (3b).

(3) a. *O Janis δjavazi ksana to vivlio*
The John reads again the book
'John reads again the book'
b. *O Janis ksanaδjavazi to vivlio*
The John again-reads the book

*Kse-* is used only as a preverb and cannot appear independently as a word-form:

(4) a. *O Janis ksekliδose tin porta*
The John unlocked the door
'John unlocked the door'
b. **O Janis kliδose kse tin porta*

The meaning of *kse-* that is used in its most productive formations is that of reversing the event (4a), but *kse-* has also an "intensive" character and can denote a high degree of realization of the verbal properties:

(5) a. *ksaγripno* vs *aγripno*
to be awake to stay up
b. *ksalafrono* *alafrono*
to relieve to lighten

With the intensive meaning, *kse-* also appears as a nominal prefix, but it will not particularly concern us here since it falls out of the scope of this study.

(6)

| | *kse-* | Noun | Adjective |
|---|---|---|---|
| a. | *ksekalokero*<br>end of summer | *kalokeri*<br>summer | |
| b. | *ksefoto*<br>place with a lot of light, surrounded by dark | *fos*<br>light | |
| c. | *ksekaθaro*<br>clearcut | | *kaθaro*<br>clean |
| d. | *kseksaspro*<br>all-white | | *aspro*<br>white |

In Ancient Greek, *para-* was used both as a preposition and as a preverb. The ancient prepositional use of *para-* is extremely limited today. It appears in fixed expressions and still denotes the idea of proximity or parallelism to the notion expressed by the item it is combined with:[6]

(7) a. *Para liyo ke θa jinotan proθipuryos*
For little and he would have become prime minister
'He almost became prime minister'

b. *Ine eksi para tetarto*
It is six to quarter
'It is a quarter to six'

Besides its limited use as a preposition, *para-* also appears attached to nouns and adjectives (8) or to verbs (9–11). As a prenominal, it keeps the idea of proximity and/or parallelism, and sometimes, develops a pejorative meaning (see (8a, b)).

| | | | | |
|---|---|---|---|---|
| (8) | *para-* | | Noun | Adjective |
| a. | *parapeδia* | vs. | *peδia* | |
| | unofficial side-education | | education | |
| | 'shadow-education' | | | |
| b. | *parakeros* | | *keros* | |
| | out of time-ADJ | | | time |
| c. | *paromios* | | | *omios* |
| | similar | | | identical |
| d. | *paraθalasios* | | | *θalasios* |
| | sea-side-ADJ | | | maritime |

*Para-* as a preverb still appears with its ancient meaning of proximity or parallelism in several verbs of Ancient Greek origin (9):

(9) Ancient Greek and Modern Greek

| | | | |
|---|---|---|---|
| a. | *paravalo* | vs | *valo* |
| | to compare | | to put |
| b. | *parangelo* | | *angelo* |
| | to order, to command | | to announce |

Notice, however, that in recent times, *para-* has also developed the meaning of an excessive realization of the notion expressed by the verb (see Triantaphyllides 1991).

(10) *I Maria parakimiθike xtes ke simera δen nistazi*
The Mary overslept yesterday and today NEG BE.3SG.sleepy
'Mary overslept yesterday and today, she is not sleepy'

Like *kse*-, the preverb, *para*- cannot be separated from the verbal base and used as an independent word:

(11) a. *Xtes paraefaje ke δen esθanete kala simera*
Yesterday he ate a lot and does not feel well today
b. **Xtes efaje para ke δen esθanete kala simera*

## 3. The semantics of *ksana, kse- and para-*

Generally, *ksana* does not bring any change to the root meaning, the argument structure, or the aktionsart of the verb it combines with.[7] The semantic interpretation of the composite containing *ksana* with a verb is usually fully compositional:[8] *ksana* provides an iteration to the event, that is an external specification to the verb. This is evidenced in (12) in which two verbs denoting an accomplishment, *vafo* 'to paint' and *ksanavafo* 'to repaint' allow frame-temporal expressions, e.g., *se pende ores* 'in five hours', but not durative adverbial expressions such as *ja pende ores* 'for five hours'.

(12) a. **(Ksana)vafi to spiti ja pende ores*
He (re)paints the house for five hours
b. *(Ksana)vafi to spiti se pende ores*
He (re)paints the house in five hours

As opposed to verbs with *ksana*, the *kse*- and *para*- verbal formations do not show a consistent behavior with respect to semantics.

Firstly, there are occurrences where *kse*- or *para*- do not affect the aspectual properties and the argument structure of the verbal base. In that sense, they behave like *ksana*, as seen in (13):

(13)

| a. Ver | b. *kse-* | c. *para-* | d. *ksana* |
|---|---|---|---|
| *kliδono* | *ksekliδno* | *parakliδono* | *ksanakliδono* |
| to lock | to unlock | to overlock | to relock |
| *vulono* | *ksevulono* | *paravulono* | *ksanavulono* |
| to block | to unblock | to overblock | to reblock |

In these examples, both the plain verb and the occurrences with *ksana, para-* and *kse*- denote an accomplishment. *ksana*, *para-*, and *kse*- provide an external specification to what is expressed by the verb, that is they reverse the event (*kse*-) or realize it excessively (*para*-).

Notice, however, that while *ksana* may combine with all kinds of verbs, that is with verbs that denote a state (15a), an activity (15b), an achievement

(15c), and an accomplishment (15d), *kse*- and *para*- are submitted to specific restrictions (16-17):

(14) Verbs used below
*kimame* ‘to sleep’, *zo* ‘to live’, *xorevo* ‘to dance’, *trexo* ‘to run’, *ftano* ‘to arrive’, *xano* ‘to lose’, *niko* ‘to win’, *laspono* ‘to cover with mud’, *skonizo* ‘to cover with dust’.

(15) *ksana*

| a. | b. | c. | d. |
|---|---|---|---|
| *ksanakimame* | *ksanaxorevo* | *ksanaftano* | *ksanalaspono* |
| to re-sleep | to re-dance | to re-arrive | to recover with mud |
| *ksanazo* | *ksanatrexo* | *ksanaxano* | *ksanaskonizo* |
| to re-live | to re-run | to re-loose | to recover with dust |
| *ksanalipame* | *ksanakolibo* | *ksananiko* | *ksanavafo* |
| to be sorry again | to re-swim | to re-win | to repaint |

(16) *kse*-

| a. | b. | c. | d. |
|---|---|---|---|
| **ksekimame* | **ksexorevo* | **kseftano* | *kselaspono* |
| | | | to scrape the mud of |
| **ksezo* | **ksetrexo* | **ksexano* | *kseskonizo* |
| | | | to dust |
| **kselipame* | **ksekolibo* | **kseniko* | *ksevafo* |
| | | | to unpaint |

(17) *para*-

| a. | b. | c. | d. |
|---|---|---|---|
| *parakimame* | *paraxorevo* | **paraftano* | *paralaspono* |
| to oversleep | to over-dance | | to over-cover with mud |
| *parazo* | *paratrexo* | **paraxano*[9] | *paraskonizo* |
| to over-live | to overrun | | to over-cover with dust |
| *paralipame* | *parakolibo* | **paraniko* | *paravafo* |
| to (over)be sorry | to over-swim | | to over-paint |

As shown by the examples in (16) and (17), *kse*- productively combines with verbs denoting an accomplishment, while *para*- may accept verbs denoting a state, an activity and an accomplishment. Both *kse* -and *para*- cannot be adjoined to verbs denoting an achievement.

Secondly, the three preverbs do not only differ with respect to the selectional restrictions that they impose to the base. As opposed to *ksana*, there are also occurrences with *kse*- and *para*- where the two preverbs do not have the same meaning as the one shown in the examples of (15–17), that is they do not reverse (*kse*-) or overdo the event (*para*-). Look at the examples in (18):

(18) a. Verb b. *kse-* c. *para-*

| a. Verb | b. *kse-* | c. *para-* |
|---|---|---|
| *kaθarizo* | *ksekaθarizo* | |
| to clean | to clear-up | |
| *elafrono* | *ksalafrono* | |
| to lighten | to relieve | |
| *amolo* | *ksamolo* | |
| to lance | to let loose, release | |
| *pulo* | *ksepulo* | |
| to sell | to sell-out | |
| *vlepo* | | *paravlepo* |
| to see | | to overlook, to neglect |
| *kino* | | *parakino* |
| to move | | to instigate |
| *γrafo* | | *paraγrafo* |
| to write | | to erase |
| *akuo* | | *parakuo* |
| to hear | | to misunderstand, to disobey |

Considering the verb to be a notional domain with a set of properties, *kse-*, in the examples of (18), contributes to bring all these properties to a high degree of realization and assumes a function of intensifying the meaning of the root verb. A similar behavior is attested with the occurrences of *para-* in (18): *para* denotes a proximity, a parallelism to the meaning of the verb, a meaning also found in Ancient Greek formations,[10] but also bears the notion of depassing what is denoted by the verb. It is under the intensive meaning that *kse-* can combine with some verbs of achievement, and form verbs with a non-compositional, and highly lexicalized meaning, something that would not have been possible if *kse-* had a reversative meaning (see (19a, b)). It is under the meaning of proximity, or parallelism to what is expressed by the verb, that *para-* may also be combined with some verbs of achievement, a combination that would have been impossible if *para-* had the meaning of excess (see (19b, c)).

(19)

| | Verb | *kse-* | *para-* |
|---|---|---|---|
| a. | *fevyo* | *ksefevyo* | |
| | to leave | to escape | |
| b. | *pefto* | *ksepefto* | *parapefto* |
| | to fall | to degrade | to fall aside, to be mislaid |
| c. | *vjeno* | | *paravjeno* |
| | to go out | | to compete |

Furthermore, if *kse-* and *para-* can affect the aspectual structure of the root verb,[11] we would also expect that the two preverbs can also alter the argument structure of the root verb.[12] In fact, the examples in (20) illustrate this phenomenon.

| | | | |
|---|---|---|---|
| (20) | a. Verb | b. *kse-* | c. *para-* |
| | *aplono* | *ksaplono* | |
| | to lay | to lie-down | |
| | *vrazo* | *ksevrazo* | |
| | to boil | to wash-up,<br>'to wash ashore, to throw-out' | |
| | *δino* | *kseδino* | |
| | to give | to relax | |
| | *jelo* | *ksejelo* | |
| | to laugh | to cheat | |
| | *δerno* | | *paraδerno* |
| | to beat | | to drift |

Following the difference in meaning and syntactic behavior, the question that arises now is whether there are instances of homophonous *kse-* and *para-*, or we deal with cases of polysemous items.

In accordance with Lieber and Baayen 1994 (hereafter L&B), I claim that each of these preverbs exhibits a polysemous meaning, since both *kse-* and *para-*, in their various uses, are more or less semantically coherent. I would like to propose that, in the lexicon, there are only two lexical entries, *kse-* and *para-*, to which a semantic variation can occur, depending on their combination with specific verbal bases.[13]

If we consider that there is a single *kse-* in the lexicon, we may tentatively suppose that

a. *kse-* has a general semantic interpretation as the item which intensifies the meaning of the root verb, and
b. *kse-* assumes the reversative meaning when its ability of inducing the verbal properties to a high degree of realization crosses the limits of the notional domain of the verb, up to a point of the complete refutation of its properties.

A similar argumentation may also be used for *para-*.

a. We saw before that *para-* expresses the idea of proximity or parallelism to the meaning of the verb.

b. It is this notion of proximity or parallelism that is pushed to an exaggeration, and may trigger an excessive realization of the verbal properties (see the examples in (17). That is why the "excessive" *para-* is used with verbs of states only when their meaning is taken to be a series of repeated processes. In this case, *para-* contributes to a confirmation of the process denoted by the verb. For instance, as seen below, the derived *parazo* 'to over-live' < *para* + *zo*'to live' cannot be used without a heavy stress and an adverbial expression denoting the situation or the conditions under which these processes have taken place.

(21) a. **O Janis parazi*
The John overlives
vs.

b.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *O* | *Janis* | *parazi* | *kala* | *meta* | *to* | *γamo* | *tu*. |
| The | John | overlives | well | after | the | marriage | his |

'John lives well indeed after his marriage'

In L&B's (1994) paper, there is a connection between the semantics of polysemous preverbs and their productivity suggesting that the more determined the semantic representation of a lexical item is, the more productive the item can be (L&B: 70). In fact, this is what we see in the case of *ksana* instances compared to *kse-* and *para-* ones. As already said, *ksana* has only one particular meaning. It expresses the iteration and can be combined with all kinds of verbs (see (15)). As such, it is an extremely productive preverb. *Kse-* and *para-*, on the other hand, are less determinate in meaning. They are not freely adjoined to all kinds of verbs, and are certainly less productive than *ksana.* Moreover, if we leave aside *ksana* and compare *kse-* and *para-*, we see that *para-* is more productive than *kse-* since *para-* can combine with more types of verbs (compare (16) and (17) above). This can also be explained by the fact that the semantic representation of *para-* is more fixed than the one of *kse-*. We saw before that *para-* expresses the idea of proximity or parallelism since Ancient Greek, while *kse-* has a general semantic interpretation as the item which provides a high degree of realization of the verbal properties.

To sum up, on the basis of the selectional restrictions and the general semantic specifications that the three preverbs bring to the verbal base, I would like to suggest that *ksana, kse-,* and *para-* can be distributed into two categories, internal and external preverbs, following a distinction proposed by Di Sciullo (1997) for the French preverbs. I assume that *ksana* is an external preverb, while *kse-* and *para-* display rather a dual character.[14] We have already seen that *ksana* does not require any particular selectional restrictions in the combination with verbs, and provides an external specification to the verbal notional domain, since it does not trigger any change to the aspectual structure or to the argument structure of the root verb. *kse-* and *para-*, however, constitute interesting cases. On the one hand (see (18)–(20)), they seem to determine the internal properties of the root verb since they affect its meaning, its aspectual properties, and its argument structure. On the other hand, in other cases, they bring only adverbial and external specifications to the verb (see (16), (17)). In the latter, they resemble to *ksana,* but contrary to it, are submitted to selectional restrictions with respect to the verbal bases they attach to, in that they do not combine with all kinds of verbs. An approach that would treat the *kse-* and *para-* cases as pure syntactic adjunction (e.g., through verb raising) or as clearcut instances of incorporation would fail to account for their peculiar behavior: their characteristics do not fulfill the basic criteria of full productivity and semantic compositionality that are proper to syntactically relevant processes. Thus, on the basis of considerations taken from semantics, there are reasons to provide a morphological account for both *kse-* and *para-*, while for *ksana* further evidence should be looked for. In fact, after testing the criteria of productivity and compositionality with 36 verbs combined with *ksana*, Smyrniotopoulos and Joseph (1997, 1998) note that, in spite of some exceptions, a syntactic treatment of the *ksana*-verb combinations may be possible since *ksana* occurs more or less freely with verbs.

## 4. Structural considerations

### 4.1 The notion of headedness

As already shown, *ksana*, *kse-*, and *para-* are productively combined with verbs to form verbs. Thus, they do not bring any categorial change to the verbal base.

(22) a. *ksanajenjeme* < *ksana* *jenjeme*
to be reborn re to be born
*ksanaδjavazo* < *ksana* *δjavazo*
to reread re to read
b. *ksefevγo* < *kse* *fevγo*
to escape to leave
*ksekliδono* < *kse* *kliδono*
to unlock un to lock

c. *parajemizo* < *para jemizo*
to overstuff over to stuff
*paraserno* < *para serno*
to beguile to drag

However, while *ksana* is always combined with a verbal base, it is possible for *kse*-and *para*- to combine with a nominal base. In the latter case, the derived word either keeps the nominal category of the base (see (6) and (8) above), or is a verb (23):

(23) a. *ksefonizo*, **fonizo*, *foni*
to shout voice
b. *ksemaljazo* **maljazo* *mali*
to ruffle, to dishevel hair
c. *parangonizo* **angonizo* *angonas*
to push aside elbow
d. *parastrato* **strato* *strata*
to stray road
e. *ksemonaxjazo* **monaxjazo* *monaxos*
to isolate alone

A closer look at the examples in (23) reveals that the string that follows *kse*- and/or *para*- is not an attested word. For instance, **fonizo* does not exist, while *ksefonizo* "to shout" (23a) is a well-attested verb. I would like to claim that strings like *foniz*- may be viewed as "bound stems", that is as stems that do not become words with the addition of appropriate inflectional endings (in our case with the addition of -o).[15] Bound stems are most common in the history of the Greek language, since the Homeric period. Notice also that in (23), the verbal category is not given by one of the preverbs, but by one of the verbal derivational suffixes -*iz*-, -*az*-, -*ev*-, -*on*-, -*o*[16] that most productively participate in Greek verbal derivation. Thus, even in cases involving a bound stem, *kse*- and *para*- could not be considered to be responsible for changing the category of the bases they are combined with.

Since no category change is involved in the combination of verbs with one of the preverbs, we may conclude that the three preverbs are not heads of their structures.[17] This conclusion is in accordance with the general property of Modern Greek to have right-headed structures in word formation (see Ralli 1988 for the notions of headedness and directionality). If heads are the rightmost categories in words (or $X^0$ expressions), and the leftmost preverbs are non-heads, I would further propose, following Di Sciullo (1997), that the preverbs are either adjuncts or complements. The preverbs do not saturate any of the arguments of the verbal heads, therefore, they can only be adjuncts in the structures they are part of.

### 4.2 The separability criterion

As already shown, *kse-* appears only in clusters/composites and never as an independent word (see (4)), while *para-* is always a bound form when used as a preverb (see (11)). Contrary to *kse-* and *para-*, however, *ksana* can combine with a root verb (24a) or may be used as an independent word-form in several positions of the sentence (25), without any significant change to the meaning of the sentence:

(24) a. *Ksanaδjavasa* to *iδjo* *vivlio* *jati* *itan* *enδiaferon*
(I) re-read the same book because (it) was interesting
vs.
b. *δjavasa* *ksana* *to* *iδjo* *vivlio* *jati* *itan* *enδiaferon*
(I) read again the same book because (it) was interesting"

(25) a. *δjavasa ksana to iδjo vivlio jati itan enδiaferon*
b. *δjavasa to iδjo vivlio ksana jati itan enδiaferon*
c. *ksana δjavasa to iδjo vivlio jati itan enδiaferon*[18]

Thus a plausible hypothesis would be that the [*ksana*-verb] composite is formed in syntax by a process such as incorporation (see Baker 1988), or verb raising (surface-structure adjunction).[19] In fact, an incorporation analysis has been postulated by Rivero (1992) who justifies her proposal by claiming that *ksana* belongs to the argument structure of the verb and structurally, can be analyzed as VP-internal. Rivero's claims about a syntactic account of *ksana* have been severely questioned by Kakouriotes, Papastathi and Tsangalidis (1997) who claim that she offers no sufficient and independent evidence for distinguishing *ksana* from other similar adverbs that fail to incorporate (e.g., *pali* 'again' vs. **paliγrafo*'write again'). In the same vein, Smyrniotopoulos and Joseph (1997: 120) note that although the *ksana*-verb formations are very productive, they do not fully respond to the following predictions that usually should hold in case of a syntactic incorporation account.

a. For every phrasal combination of Verb +Adverb, there exists a corresponding composite. The dubious acceptability of the verb *?*ksanaperijelo* 're-mock' (< *ksana* + *perijelo*'mock'), as opposed to the perfectly acceptable phrasal form *ton perijelasan ksana* 'they mocked him again', constitutes an exception to this prediction (S&J 1998: 456).

b. If there is no phrasal combination, there is no corresponding composite and every composite has a phrasal counterpart. It should be noticed that to the composite *ksananjono* 'to rejuvenate, to become young again', mentioned by Mendez-Dosuna (1997), there is no independent phrase *njono ksana* or independent verb *njono.*[20]

c. Every composite is compositional in meaning, and shows no idiosyncratic meaning differences from its phrasal source. However, there are *ksana*-verb composites that develop a non-predictable meaning (as well as a non-predictable syntactic behavior) that is not determined compositionally from the combination of *ksana* with the verb. Consider the examples in (26) for an illustration to this last observation.

(26) a. *δen prosekse ke ksanakilise*
(S/he) wasn't careful and relapsed
ksanapjanete me tin iδja δulja
(S/he) is re-taken with the same job
'She starts again the same job'

b. **δen prosekse ke kilise ksana*
*δen pjanete ksana me tin iδja δulja*
vs.

c. *δen prosekse ke kilise ksana sto vurko*
(S/he) wasn't careful and scrolled again in the mud
*δen pjanete ksana stin pajiδa*
(S/he) is taken again in the trap

In addition, it should be noticed that there are also verbs that do not clearly accept *ksana* in the preverbal position, while they accept *ksana* as a separable modifier in the same context:

(27) a. *Ise ksana stin iδja δulja*?
Are (you) again in the same job?

b. *Exis ksana tin iδja δulja*?
Have (you) again the same job?
But

(28) a. ??*Ksanaise stin iδja δulja*?
Re-are (you) in the same job?

b. ??*Ksanaexis tin iδja δulja*?
Re-have (you) the same job?

To sum up, on the basis of the separability criterion, there is little doubt that *kse*-and *para*- are bound morphemes. As such, they could be given the prefixal status, while for *ksana* the situation is not so clear and further evidence should be looked for. On the one hand, *ksana* may be separated from the verbal base and used as an independent word in several positions of the sentence (see (24) and (25)). As such, it shares the properties with adverbs. On the other hand, as pointed out by Smyrniotopoulos and Joseph (1997, 1998), and further illustrated in (26-28), some *ksana* formations exhibit a behavior that is close to *kse-* and *para-* formations.

### 4.3 Co-occurrence and alternation between preverbs

As expected, very productive preverbs may co-occur and alternate between them. In fact, this is usually attested with occurrences containing *para-* and *ksana:*

| | | | | |
|---|---|---|---|---|
| (29) | a. | *parafuskono* | *ksanaparafuskono* | *fuskono* |
| | | to re-inflate | to over-re-inflate | to inflate |
| | b. | *ksanafuskono* | *paraksanafuskono* | |
| | | to re-inflate | to over-re-inflate | |
| | c. | *paratendono* | *ksanaparatendono* | *tendono* |
| | | to over-stretch | to re-over-stretch | to stretch |
| | d. | *ksanatendono* | *paraksanatendono* | |
| | | to re-stretch | to over-re-stretch | |

Notice, however, that an alternation between the two preverbs is possible only when *para-* is used as external. When *para-* is internal, the ordering between the two preverbs imposes the presence of *para-* to be close to the verb, while *ksana* is added outside the cluster [*para*-verb]:

| | | | | | |
|---|---|---|---|---|---|
| (30) | a. | *parakino* | *ksanaparakino* | **paraksanakino* | *kino* |
| | | to instigate | to re-instigate | | to move |
| | b. | *paravlepo* | *ksanaparavlepo* | **paraksanavlepo* | *vlepo* |
| | | to overlook | to re-overlook | | to see |

With respect to *kse-*, it is crucial that there is no alternation when both *kse-* and *para-* are external, or between external *kse-* and *ksana* (31a). External *kse-* is always closer to the verbal root, while *para-* and *ksana* appear at the periphery. This strict ordering between *kse-*, *ksana* and external *para-* also holds when *kse-* assumes an internal role, that is when it has the intensive meaning (31b).

| | | | | |
|---|---|---|---|---|
| (31) | a. | *fuskono* | b. | *kaθarizo* |
| | | to inflate | | to clean |
| | | *ksefuskono* | | *ksekaθarizo* |

| | |
|---|---|
| to deflate | to clear-up |
| *paraksefuskono* | *parakseka θarizo* |
| to over-deflate | to over-clear-up |
| *ksanaksefuskono* | *ksanakseka θarizo* |
| to re-deflate | to re-clear-up |
| *ksanaparaksefuskono* | *ksanaparakseka θarizo* |
| to re-over-deflate | to re-over-clear-up |
| *paraksanaksefuskono* | *paraksanakseka θarizo* |
| to over-re-deflate | to over-re-clear-up |
| **kseparafuskono* | **kseparaka θarizo* |
| **kseksanafuskono* | **kseksanaka θarizo* |

Contrary to the examples of (31) that show a strict ordering between external *kse-* and *para-*, there is usually a mutual exclusion between the two in their internal use (see 32a–e).

(32) a. *kino*
to move
b. *ksekino*
to start-up, to move-off
c. *parakino*
to instigate
d. **kseparakino*
e. **paraksekino*

n (33), the external preverbs *ksana* and *para-* allow iteration while the external *kse-* that adds an endpoint to the meaning of the verb shows no iteration.

| | | | |
|---|---|---|---|
| (33) a. | *kaθarizo* | *ksanakaθarizo* | *ksanaksanakaθarizo* |
| | to clean | to re-clean | to re-re-clean |
| b. | | *parakaθarizo* | *paraparakaθarizo* |
| | | to over-clean | to over-over-clean |
| c. | | *ksekaθarizo* | **kseksekaθarizo* |
| | | to clear-up | |

Finally, as shown in (34), an iteration of *para-* is possible when the most left-handed occurrence acts as an external prefix and the other occurrence as an internal one (34d), but not in the opposite order (34e).

| (34) a. verb | b. internal *para* | c. external *para* | d. $\text{para}_{ext}$-$\text{para}_{int}$ | e. $\text{para}_{int}$-$\text{para}_{ext}$ |
|---|---|---|---|---|
| *vlepo* | *paravlepo* | *paravlepo* | *paraparavlepo* | **paraparavlepo* |
| to see | to overlook | to over-see | to over-overlook | |

## 5. Phonological evidence

Evidence taken from phonology supports the prefixal status of *kse*- and *para*-. In Greek nominal prefixation, there is usually a stress shift to the antepenultimate syllable. For example, the adjective '*aγnostos*'unknown' is stressed on the prefix *a*- 'un', while the adjective *γno'stos* 'known' bears a stress on the last syllable. As claimed by Nespor & Ralli (1996), the stress shift is caused by the prefixation of *a*- to *γnostos*. If we look at the formations of (6) and (8), we observe that the same stress-shift phenomenon occurs in words like *ksekalokero*, *parakeros*, etc., which I repeat here for convenience:

| | | | | |
|---|---|---|---|---|
| (35) | Noun | | *kse*- | *para*- |
| | a. | *kalo'keri* | *kseka'lokero* | |
| | | summer | end of summer | |
| | b. | *ke'ros* | | *pa'rakeros* |
| | | time | | out of time |

Stress-shift is not a safe criterion for deciding about the prefixal status of *ksana*, since *ksana* appears only with verbs, and prefixed verbs are not generally submitted to a stress-shift procedure. However, it is crucial that all [*ksana*-verb] clusters bear only one stress, that is they constitute one phonological word, while *ksana* and the root verb constitute phonological words on their own if they are taken separately. Since in Greek the boundaries between structural words and phonological words generally coincide, we may, thus, assume that the *ksana* formations display word-like properties.

It should be noticed that when a preverb is attached to a verb root beginning by a vowel, a vowel deletion may occur at the morpheme boundary between the two. As the following examples illustrate, this vowel deletion is not obligatory for all preverbs though.

| | | | | | |
|---|---|---|---|---|---|
| (36) | | Verb | *kse*- | *para*- | *ksana* |
| | a. | *exo* | *ksexo* /**kseexo* | *parexo* /**paraexo* | *ksanaxo* / *ksanaexo*[21] |
| | | [*kse*+*exo*] | | [*para*+*exo*] | [*ksana*+*exo*] |
| | | to have t | o un-have | to provide | to re-have |
| | | | | **parexo* / *paraexo* | |
| | | | | to over-have | |
| | b. | *aniyo* | *ksaniyo* /**kseaniyo* | *paraniyo*/ *paraaniyo* | *ksananiyo* / *ksanaaniyo* |
| | | | [*kse*+*anigo*] | [*para*+*anigo*] | [*ksana*+*anigo*] |
| | | to open | to open-out | to over-open | to re-open |

In (36), a vowel deletion always occurs in case of *kse-*, in both its internal and external use, and optionally occurs in case of *ksana*. As for the formations with *para-*, they are subject to an optional vowel deletion when *para-* functions as an external preverb and to an obligatory vowel deletion when *para-* has an internal character. *kse-* and internal *para-* seem to be more lexically bound with the verb root since an obligatory vowel deletion occurs at the boundary between the two.[22] Given the fact that no obligatory vowel deletion occurs in case of external *para-* and *ksana*, they seem to have a more loose structural relation with the verbal base, although it might be the case that the internal/external distinction does not hold for morpho-phonological structures. Thus, a possible distinction between prefixes and non-prefixes does not capture the peculiarity of vowel deletion. It follows that a more refined account of the various differences and similarities displayed by the three preverbs is necessary.

## 6. Partial conclusions

Following evidence from semantics, the separability criterion, and further evidence taken from phonology, we have already concluded that *kse-* and *para-* are prefixes. As such, they should be subject to a word-formation analysis. Evidence from the co-occurrence and alternation between preverbs shows that *ksana* freely alternates with the external *para* (see examples in (29)). This free alternation leads us to suppose that both the *para-* and *ksana* formations should be accounted for in a similar manner. That is, if prefixation is to be treated as a morphological procedure, the *ksana* formations should receive a morphological account too. These observations do not force us to consider the formations of [*ksana*–Verb] as cases of prefixation, but they do show that these formations cannot be a matter of syntax.[23] This would explain why some *ksana* formations are peculiar with respect to their meaning, their structural behavior, as shown in (26–28) above, and their word-like stressing properties. A morphological account of *ksana* would raise the question, however, of whether *ksana* is a prefix or a word participating in compound formations.[24] Since *ksana* is not a bound morpheme and can appear independently in sentences (see (24–25), a plausible answer would be that *ksana* is a word form that actively participates in compound formations. This solution would accommodate the following issues:

a. it would provide a unified account for the three preverbs, *kse-*, *para-* and *ksana*, that is a morphological account, since prefixation (cases of *kse-* and *para-)* and compounding (*ksana*) are clearly morphological processes in Greek, as shown by Ralli 1988 and 1992.

b. It would allow us to differentiate between *ksana* on the one hand and *kse-* and *para-* on the other, since prefixes are more bound with the stem bases on several aspects, phonological, structural, and semantic, while compounding displays a greater transparency in both form and meaning. As shown above, *kse-* and *para-* can impose selectional restrictions on the verbal root, while *ksana* combines more or less freely with all kinds of verbs (see (15)).

It should be noticed that the attested differences in the co-occurrence and alternation between *kse-* and *para-,* as well as with respect to vowel deletion, suggest that the two preverbs should also be distinguished. As seen in (31a), although the reversative *kse-* may act as an external preverb, it appears to be the closest to the verbal root since it does not alternate with external *para-* and *ksana. Para-* though displays a more differentiated character. With the external function (meaning of exaggeration and excessive realization of the verbal properties), it alternates with *ksana* (29), and appears outside the [*kse*-Verb] clusters (31a). With the internal function (meaning of proximity/parallelism, and/or that of depassing the verbal properties), *para-* and *kse-* are mutually exclusive (32a–e). Furthermore, as seen in Section 4, in the *para-* formations, an obligatory vowel deletion occurs when an internal *para-* is involved, as opposed to all *kse-* formations (i.e., involving either an internal *kse-* or an external one) that are always subject to vowel deletion in the appropriate context. Assuming that the three levels of linguistic analysis, i.e., phonology, morphology, and syntax are dealt with altogether,[25] a simple distinction between internal and external preverbs is not sufficient to account for these differences. In the following section, we will see how the distinct behavior of *kse-* and *para-* is taken into account in a unified morphological analysis that can also accommodate *ksana*.

## 7. Theoretical analysis

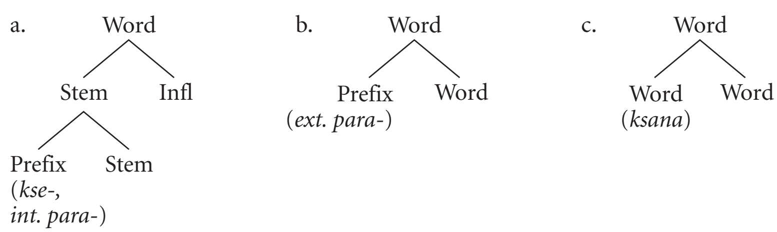
Structurally, *kse-* and *para-* behave like prefixes, while *ksana* is a word, since it appears independently in syntactic formations. As such, *kse-* and *para-* participate in the derivational process of prefixation, while *ksana* takes part in compounding.[26] Phonological evidence has shown that the structural relation between *kse-* and internal *para-* is more bound than the one involving *ksana* and external *para-*, since the latter are not subject to an obligatory vowel deletion in the appropriate context. Finally, no category change is involved in the structures where the preverbs participate. Assuming that Greek word formation is right-headed, *ksana*, *kse-*, and *para-* are thus adjuncts in their structures. Generally, adjuncts can add information without changing the syntactic category of the configuration they are part of, and without serving as an argument to the head.[27]

We will see now how a morphological account takes into consideration the relevant structural, phonological and semantic differences between prefixes on the one hand, and between prefixes and simple compound members on the other.

The basic assumptions of this approach, as presented in Ralli (1999a, 2001), are the following:

a. Morphology is a grammatical module[28] generating morphological expressions in the computational space of the faculty of language.
b. In an inflectionally rich language, e.g., Greek, a morpheme-based lexicon feeds both Morphology and Syntax, where entries may be words ($X^0$), but also units smaller than words (i.e., stems).
c. Morphology is responsible for constructing binary well-formed structures in a sequence of steps, relating heads and non-heads, that is adjuncts or complements with heads.[29] Non-heads with the role of complements are usually found in synthetic/verbal compound structures, where the non-head member saturates one of the arguments of the verbal head.[30] The rest of non-heads function as adjuncts.

Within the spirit of these assumptions, I would like to propose that prefixed and compound words involving the addition of *ksana, kse-* and *para-* have the structures as in (37), where the basic morphological categories are those of Word, Stem, Prefix, and Inflectional Suffix.[31]

(37)



In (37a), a prefix is added to a stem to build a prefixed stem that becomes a word after the addition of the appropriate inflectional affix. I claim that this is the case of all *kse*- prefixation (internal and external) and the less productive part of *para*- prefixation, the one involving internal *para*- (see examples in (18) and (19)). In (37b) and (37c), a prefix, or a word, are added to another word to form a prefixed or a compound word respectively. These structures generate the *ksana* compounds (37c) or the most productive external *para*- formations (37b). In other words, the distribution of adjunction sites is asymmetrical with respect to the distribution of prefixes in internal and external ones: while *kse*- is added only to stems, and *ksana* only to words, there are two possible adjunction sites for *para*-, dependently on its particular meaning and structural behavior seen so far. [32]

It is crucial that the distinction between stem adjunction and word adjunction for the study of the three preverbs accounts for both the similarities and the differences between them.

More specifically, the configurational difference between the structures generating the three preverbs, as presented in (37), correctly predicts the following characteristics:

a. Prefixes of the same semantically-defined category (e.g., external) may display a different structural behavior since they may have different adjunction sites in word formation. For instance, we saw that although *kse*- can be external (i.e., with a reversative meaning), it does not alternate with another external prefix, such as *para*-.
b. A word-based prefix must precede a stem-based one. Data exposed in the section of the co-occurrence and alternation of the preverbs show that this is exactly the case with the preverbs under consideration. *Ksana* and the most productive part of *para* precede the internal *kse*- as well as the less productive part of *para*-, while they freely alternate between them.
c. The difference in adjunction sites cuts across the difference between bound and free morphemes (e.g., between prefixes and words). For instance, we saw that although the external *para*- is a prefix, it shares more similarities with the non-prefix *ksana* than with the prefix *kse*-.
d. In a morpheme-based language like Greek, the possibility to have adjunction to words allows us to account for cases of lexical insertion in both morphology and syntax, involving the same lexical item. For example, a word like *ksana* has an active role in syntax (see (25)), but this does not prohibit it from participating in the process of compounding in morphology. Adjunction to stems in word formation predicts that the morphological structure could involve constituents that may not appear as actual words after the addition of the appropriate inflectional affix. We saw above

that *kse*- and internal *para*- are prefixes that participate in derivational processes where the verbal part following the prefix is not an actual word. This has been illustrated in (23). It is important to note that nominals containing the word-based *para*- and *ksana* do not allow similar formations, since a word-based adjunction would require that the word that functions as the base of the formation is part of the existing words of the language.

e. Finally, the different adjunction sites for *ksana, kse*- and *para*- also predict that there must be some phonological differences in the items involving the combination of a preverb with a verb. In fact, as illustrated in (36), in a stem-based structure involving all instances of *kse*- and *para*-, an obligatory vowel deletion occurs if the verbal base begins by a vowel. On the contrary, in a correspondent word-based structure involving external *para*- and *ksana* vowel deletion has no obligatory character.

## 8. Summary

In this work, three productive Greek preverbs have been examined, namely *ksana, kse*- and *para*-. On the basis of their semantic interpretation, structural characteristics and phonological behavior in the structures they are part of, it was shown that the three preverbs are distinguished into the following categories:

a. Internal and external preverbs, according to the specifications that they bring to the verbal base, that is with respect to the root meaning, the aspectual structure or the argument structure of the verb. *Kse*- and *para*- display a dual character, since they can assume an internal or an external role, while *ksana* is used only as an external preverb.
b. Prefixes and non-prefixes (words), depending on the ability to appear as bound or independent elements in words and sentences respectively, and according to phonological phenomena that are triggered when a preverb is combined with a verbal base. While there is no doubt about the prefixal status of *kse*- and *para-, ksana* is rather a word, since it can be separated from the verbal base without any change to the meaning of the sentence it is part of. It was claimed that *ksana* actively participates in compound word-formations, while *kse*- and *para*- are handled as cases of derivational prefixation.

It was shown, however, that this subcategorization of the three preverbs does not take into consideration all differences and similarities between the preverbs and that there are some properties that cut across these categories. In order to account for the peculiar behavior of the three preverbs, it was proposed that the combination of each preverb with a verbal base may occur in different adjunction sites within morphology. *Kse-* that appears to be closest to the base is generated as an adjunct to stems, *ksana* that has a loose relation with the base is an adjunct to words, while *para-* can be both as a stem or a word adjunct depending on the case.

An approach that handles the three preverbs within morphology, and explains their characteristics in terms of different configurations in word formation, seems to provide a better account than an approach that would have ignored morphology.

## Notes

* Parts of this paper have been presented at the workshop on Preverbs (University of Nijmegen: Jan. 19–20, 2001), at the 15th International Symposium on Greek and English (University of Thessaloniki: May 4–6, 2001), and at the Asymmetry Conference (UQAM May 7–9, 2001). I would like to thank the audiences of the three conferences for their most constructive observations. I also wish to thank G. Giannoulopoulou, D. Melissaropoulou, and Th. Tsiamas for their significant help in the location of the data that is used in this paper. I am particularly indebted to D. Cheila-Markopoulou, D. Theophanopoulou-Kontou, M. Tzevelekou, and G. Xydopoulos whose precious comments made this paper benefit greatly.

**1.** It should be noticed that these properties have been discussed for the preverbs of other languages too, such as Romance (Di Sciullo 1997, 1999) and English (Keyser & Roeper 1992).

**2.** According to Gardikas (1924) and Hatzidakis (1926), *ksana* has been formed from the combination of the Ancient Greek prepositions *ek* and *ana* (*ek* + *ana* > *eksana* > *ksana*), while *kse-* follows from the combination of the Ancient Greek preposition *ek*, denoting the origin, with the verbal syllabic augment *e* – (*ek* + *e* > *eks* + *e* > *kse*). According to Karantzola & Giannoulopoulou (2000: 194–200), there is still a semantic relation between the new formations and the Ancient Greek prepositions where they come from. The idea of repetition expressed by the ancient *ana* is present in *ksana*, and the idea of direction, or moving away from inside to outside, expressed by *ek* is indirectly present in *kse-*. For example, in the text of *Palea ke Nea Diathiki* "Old and New Testament" by Ioannikios Kartanos (issued in Venice in 1536), the verb *ksananeono* (<*eks* + *ana* + *neono*) means the reappearance of youth, while the verb *ksevlasteno* (< *ek* + *vlastano*) means the growth of buds from inside to outside of the plant.

**3.** Verbs characterized as [learned] are those that come from Ancient Greek, or constitute formations of the so-called "katharevousa", an artificial, ancient-looking form of language that was developed for political reasons by a group of literary people in the XIXth century.

4. Modern Greek examples are given in a broad phonetic transcription, while Ancient Greek ones are given in alphabetical characters.

5. *δiδo* derives from the Ancient Greek verb δiδωμι, after the loss of the ancient ending –μι.*δino* is the commonly used form in Modern Greek.

6. In this work, I restrict my attention to the use of *para-* as a preverb. In my opinion, *para-* used as a preposition (iia), adverb (iib), or conjunction (iiic) to express an opposition (meanings of 'despite' and 'but'), or an opposition in the process of comparison (meaning of 'than'), is synchronically a different homophonous item and will not concern us here.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ii. | a. | *Para* to | δjavazma | tu | δen | perase | tis | eksetasis | | |
| | | In spite of | study | his, | he didn't | pass | the | exams | | |
| | | "In spite of his study, he didn't pass the exams" | | | | | | | | |
| | b. | δen | ine | ola | ta | katastimata | anixta | *para* | mono | ta farmakia |
| | | There | are not | all | the | stores | open, | but | only | the pharmacies |
| | c. | Ine | xirotero | na les | ti γnomi su | *para* | na min ti les | | | |
| | | It is | worse | to say | your opinion | than | not to say it | | | |

However, Poulopoulou (1996) assumes a different position. Following an analysis based on pragmatics, and with the use of arguments drawn from general discourse considerations, she considers all the *para-* occurrences as different realizations of the same item. She admits though (p. 86) that there are serious difficulties in this approach.*Para* in the process of comparison has also been studied by Cheila-Markopoulou (1986).

7. See (12–15) for relevant examples. See also Smyrniotopoulos & Joseph (1997, 1998) and Kakouriotes, Papastathi & Tsangalidis (1997) for having already noticed this property.

8. See, however, Section 3.2. (26–28) for exceptions to this observation.

9. In these examples, the verb *xano* 'to lose' is taken as intransitive and not with its transitive use as in the phrase *xano xrimata* 'to lose money' where it denotes an accomplishment.

10. See Humbert (1970: 340).

11. For a solid argumentation in favor of the hypothesis that prefixes can be aspectual modifiers see Di Sciullo (1997).

12. There are two major issues which I only tackle here:

a. how arguments of verbs with preverbs are distributed among the word-internal constituents (see Neeleman & Schipper 1993 for the thematic roles in Dutch prefixation), and
b. which verbs could accept *kse-* and *para-*, contrary to the selectional restrictions illustrated in (15-17) in this paper.

Both issues need extensive research in Greek lexical semantics and I leave them to a future study.

13. That we may deal with only one *kse-* is also proposed by Efthimiou (2000) and Karantzola & Giannoulopoulou (2000: 201).

14. See Bisetto, Mutarello and Scalise (1989) for the Italian preverbs which have the same form but display various semantic characteristics.

**15.** See Ralli (1988) for a detailed account of "bound stems" in Greek morphology. See also Section 4 (36) for further details on these formations.

**16.** In the case of *parastrato* (23d), the *-o* ending is stressed and assumes the role of a derivational affix. As shown in Ralli (forthcoming), one of the productive derivational processes in Greek is the one that forms verbs out of nouns with the addition of a stressed *-o.*

**17.** Notice that sometimes a category change may occur after the combination of *kse-* and *para-* with nouns, which results into the creation of adjectives or adverbs (see the examples below). I consider these cases as instances of conversion through zero-suffixation, in the way that Neeleman & Schipper (1993) propose for the Dutch preverbs.

iv. a. *parakeros* < *para keros*
out-of-time.ADJ time, weather.N

b. *paranomos* < *para nomos*
outlaw.ADJ/N law.N

c. *ksexora* < *kse xor(os)*
out-of-place.ADV place
'independently'

**18.** In the context of the last example, *ksana* carries a heavy stress.

**19.** As Booij (1991:53–59) correctly points out for similar structures in Dutch, a deep-structure adjunction should be excluded because it would require that the verb is optionally or obligatorily subcategorized for *ksana.* Note that Booij adopts a lexical analysis for the Dutch preverbs, although he considers them as parts of phrases created in the lexicon (pp. 59-61).

**20.** *Njono* may occur in some dialects with the meaning of 'feel'.

**21.** According to Kaisse (1985), the vowel deletion in Greek is subject to the restrictions imposed by a vowel hierarchy, according to which, at the contact between two vowels, the stronger [a] triggers the deletion of the less strong [e].

**22.** A similar behavior with respect to word-internal phonology is also attested when a preverb is combined with a verbal form preceded by the augment *e-* which marks the past tense:

v. Verb int. *kse-* ext. *kse-* int. *para-* ext. *para-* *ksana*

a. *pefto ksepese/ parepese / parapese / ksanapese/*
**kseepese *paraepese ksanaepese*
[*kse+e+pese*] [*para+e+pese*] [*ksana+e+pese*]
to fall (s)he degraded (s)he fell aside (s)he re-fell

b. *dino ksedise / paraedise / ksanaedise/*
**kseedise paradise ksanadise*
[*kse+e+dise*] [*para+edise*] [*ksana+e+dise*]
to dress (s)he undressed (s)he overdressed (s)he re-dressed.

**23.** On the basis of the fact that *ksana* is recurrent in many words, Philippaki-Warburton (1970) has analyzed it as a prefix. In Xydopoulos (1996) and Kakouriotes et al. (1997), however, the *ksana* formations have the lexical status, but it is not clear whether the preverbs are analyzed as prefixes or compound constituents.

**24.** The same question is raised in Smyrniotopoulos and Joseph (1998: 481–482). Comparing *para-* with *ksana*, the authors propose a prefixal analysis for the first, while for the second they seem to be in favor of a compound status.

**25.** See note 27 for the structural differences in the derivation of a morphological expression and a syntactic one.

**26.** It should be noticed that the word status of *ksana* prevents it from combining with a linking vowel -o- in compound formations. As noted by Ralli (1999b), -o- appears between the members of a compound when the first member is a stem (e.g., *tirosalata* 'cheese salad' < *tir-o-salata*).

**27.** Following a minimalist framework, Di Sciullo (1997: 58) claims that in an adjunction structure, adjuncts do not satisfy any obligatory selectional features of the head and no checking occurs between the features of the adjunct and those of the head. Di Sciullo (1997: 57) proposes the Adjunct Identification Principle which accounts for the identification of underspecified features of heads. However, as will be clear below, the configurational analysis that I use in this paper is not based on the syntactic minimalist structure involving the structural constituents of Specifier, Complement, Head and Adjunct. In this paper, the terms of "adjunct" and "head" are used only in the general sense, that is as basic components of a general binary structural relation involving the morphological combination of two constituents, the basic of which is the head and the modifying one acts like an adjunct.

**28.** About morphology seen as an autonomous level of grammar, see also Aronoff (1994: 63) who claims that morphology is not entirely reducible to another level, and follows principles of its own, in addition to other principles that may apply to other levels as well.

**29.** See Di Sciullo (1996) on the exclusion of Specifiers in morphological structures.

**30.** See Di Sciullo and Ralli (1999) for further details on argument saturation within synthetic compounds.

**31.** These structures are generally motivated on further empirical grounds, that is on the basis of morphological data taken from Greek. For instance, it is important to note that Greek verbal and nominal categories are generally analyzed into a stem and an inflectional affix (see the upper part of the structure in (35a). All inflection and most derivation (prefixation and suffixation) are stem based, that is they involve the combination of a stem and an affix. Compounding has two kinds of structures, both of them equally productive, that is stem-based and word-based structures, where the head may be a stem or a word respectively (see Ralli 1992 and Nespor & Ralli 1996 for further details).

**32.** Notice that the idea of different adjunction sites is found in Di Sciullo's (1997) work. She proposes that prefixes are adjoined outside or inside the verbal projection, depending on their internal or external status. Although Di Sciullo adopts a morphological approach, the configurations she proposes are expressed in syntactic terms. In particular, in Di Sciullo (1996), the constituents of specifier, head, and complement are part of the XP derivation of morphological objects (where the relevant internal/external prefixation is set). However, specifiers are not part of $X^0$ head-adjunction structures which are interpreted at Morphological Form. Another significant difference between Di Sciullo's approach and the one adopted in this paper, is that in the first, the distinction between the internal and the external status is translated into a configurational difference, while in the latter there is no 1:1 correspondence between an internal/external categorization and a difference in adjunction sites which, in this paper, are expressed in morphological terms.

## References

Aronoff, M. (1994). *Morphology by Itself.* Cambridge, Mass: MIT Press.

Baker, M. (1988). *Incorporation. A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Bisetto, A., Mutarello, R. & Scalise, S. (1989). Prefissi e teoria morfologica. *Parallela, 4*, 29–41.

Booij, G. (1991). The boundary between morphology and syntax: se*para*ble complex verbs in Dutch. In G. Booij and J. Van Marle (Eds.), *Yearbook of Morphology 1990* (pp. 29–44). Dordrecht: Kluwer.

Cheila-Markopoulou, D. (1986). Ta singritika tis Neas Elinikis. Sintaktiki analisi tu singritiku vathmu ton epitheton ke epirimaton. (Comparatives of modern Greek. A syntactic analysis of the comparative degree of adjectives and adverbs). Ph.D. dissertation. Athens: University of Athens.

Di Sciullo, A. M. (1996). Atomicity and relatedness in configurational morphology. In A. M. Di Sciullo (Ed.), *Configurations. Essays on Structure and Interpretation* (pp. 17–41). Somerville: Cascadilla Press.

Di Sciullo, A. M. (1997). Prefixed-verbs and adjunct identification. In A. M. Di Sciullo (Ed.), *Projections and Interface Conditions* (pp. 52–73). New York: Oxford University Press.

Di Sciullo, A. M. (1999). Verbal structure and variation. In E. Trevino and J. Lema (Eds.), *Semantic Issues in Romance Syntax* (pp. 39–57). Amsterdam/Philadelphia: John Benjamins.

Di Sciullo, A. M. & Ralli, A. (1999). Theta-role saturation in Greek compounds. In A. Alexiadou, G. Horrocks and M. Stavrou (Eds.), *Studies in Greek Syntax* (pp. 185–200). Dordrecht: Kluwer.

Efthimiou, A. (2000). To Neoeliniko prothima kse-: I simasies tis apomakrinsis ke alajis katastasis. (The Neo-Hellenic prefix *kse-*: the concepts of move away and change of situation). *Proceedings of the 4th International Conference of Greek Linguistics* (pp. 122–130). Nicosia: University Studio Press.

Gardikas, G. (1924). The prepositions in compounding. *Athina, 24*, 75–175.

Hadzidakis, G. (1926). About the use of the preposition ek, eks (*kse-*) in Modern Greek. *Athina, 26*, 8–43.

Humbert, J. (1972). *Syntaxe Grecque*. Paris: Kincksieck. 3rd ed.

Kaisse, H. (1985). *Connected Speech. The Interaction of Syntax and Phonology.* New York: Academic Press.

Kakouriotes, A., Papastathi, M., & Tsangalidis, A. (1997). Incorporation in Modern Greek: lexical or syntactic? In Drachman et al. (Eds.), *Greek Linguistics '95. Proceedings of the 2nd International Conference on Greek Linguistics* (pp. 77–86).

Karantzola, E. & Giannoulopoulou, G. (2000). Semantic elements in compounding and derivation in Medieval Greek. *Studies of the Greek Language*, 193–202. Thessaloniki: Kyriakides.

Keyser, S. J. & Roeper, T. (1992). Re: The abstract clitic hypothesis. *Linguistic Inquiry, 23*, 89–127.

Lieber, R. & Baayen, H. (1994). Verbal prefixes in Dutch: a study in lexical conceptual structure. In G. Booij and J. Van Marle (Eds.), *Yearbook of Morphology 1993* (pp. 51–79). Dordrecht: Kluwer.

Mendez Dosuna, S. (1997). Fusion, fission, and relevance in language change: de-univerbation in Greek verb morphology. *Studies in Language, 21*, 577–612.

Neeleman, A. & Schipper, J. (1993). Verbal prefixation in Dutch: thematic evidence for conversion. In G. Booij and J. Van Marle (Eds.), *Yearbook of Morphology, 1992* (pp. 57–92). Dordrecht: Kluwer.

Nespor, M. & Ralli, A. (1996). Morphology-phonology interface: phonological domains in Greek compounds. *The Linguistic Review, 13*, 357–382.

Philippaki-Warburton, I. (1970). *On the Verb in Modern Greek*. Bloomington: Indiana University Publications.

Poulopoulou, M. (1996). Towards to a uniform semantic description of the morpheme para. M.A thesis. Crete: University of Crete.

Ralli, A. (1988). Éléments de la morphologie du Grec moderne. Ph.D. dissertation. Montreal: University of Montreal.

Ralli, A. (1992). Compounds in Modern Greek. *Rivista di Linguistica, 4* (1), 143–174.

Ralli, A. (1999a). Inflectional features and the morphological module hypothesis. *Working Papers of English and Greek Linguistics, 1*, 111–141.

Ralli, A. (1999b). To fenomeno tis sinthesis sti Nea Elliniki (The phenomenon of compounding in Modern Greek). *Parousia, IA-IB*, 183–205.

Ralli, A. (Forthcoming). *Morphology.* Athens: Patakis.

Rivero, M. L. (1992). Adverb incorporation and the syntax of adverbs in Modern Greek. *Linguistics and Philosophy, 15*, 289–331.

Smyrniotopoulos, J. & Joseph, B. D. (1997). On so-called 'adverb-incorporation' in Modern Greek. In Drachman et al. (Eds.), *Proceedings of the 2nd International Conference on Greek Linguistics* (pp. 117–128). Graz: Neubauer Verlag.

Smyrniotopoulos, J. & Joseph, B. D. (1998). Syntax versus the lexicon: incorporation and compounding in Modern Greek. *Journal of Linguistics, 34*, 447–488.

Triantaphyllides, M. (1991). *Neoelliniki Grammatiki* (A grammar of Modern Greek). Thessaloniki. 3rd ed.

Xydopoulos, G. (1996). Tense, aspect and adverbials in Modern Greek. Ph.D. dissertation. London: University College London.

# Asymmetry, syntactic objects and the Mirror Generalization

Réjean Canac Marquis
Simon Fraser University[1]

## Introduction

Since its early days, X-bar theory represents a descriptive typology of phrase markers. In its minimalist approach, Chomsky (1994 BPS, 1995) proposes to reduce fundamental aspects of X-bar theory to bare essentiels and in the process, seeks to provide a more explanatory typology of syntactic objects, i.e. phrase markers resulting from the projection of lexical items through the Merger and Move operations in the computational system (CS). In this paper, we agree with the value of deriving a typology of syntactic objects from minimalist assumptions, but argue that Chomsky's specific proposals fall short in several important respects that we discuss in details. Chomsky maintains that Xmax and Xmin projections are non-primitive, relational categories which he defines in informal terms, i.e. «. . .a projection that doesn't project any further is an Xmax, and one that is not a projection at all is an Xmin; any other is an X', invisible at the interface for computation. » (ibid: 242). However these informal definitions do not stem from any fundamental notions or conceptual necessity, but merely from minimal assumptions and do not play any further significant role in defining the whole set of possible syntactic objects.

The alternative proposal developed in this paper derives the basic distinction in syntactic projections from the fundamental role played by asymmetry in grammar, in the sense of Di Sciullo's Asymmetry Hypothesis (1999), a proposal pushing the role of asymmetry beyond the Linear Correspondance Axiom of Kayne (1994). According to Di Sciullo's hypothesis, all grammatical and structural relations in grammar must be asymmetrical, i.e. asymmetry plays a fundamental role in defining the architecture of grammars. Our main claim is that the entire typology of syntactic objects stems from a distinction between Xmax and Xnon-max projections (not equivalent to the Xmax/Xmin distinction in BPS) which itself derives from systematically computing all the primitive asymmetrical structural relations of dominance and immediate dominance between projections in a given phrase marker. In other words, the Xmax/Xnon-max distinction is not a mere assumption, but a necessary consequence of the hypothesis that asymetrical structural relations play a fundamental role in the CS. The scope of the proposal extends beyond deriving the mere Xmax/Xnon-max distinction, to predict fundamental aspects of the typology of syntactic objects.

The proposal is advantageous in several respects. Empirically, it is exhaustive in covering all possible case scenarios of Xmax and Xnon-max adjunction and substitution, contrary to BPS. It also advantageously subsumes stipulative aspects of Chomsky's system, e.g. it derives from a single source the fundamental generalization that the target of Move always projects (cf. unrelated conditions in BPS) except in exactly one case which gives rise to VP shells. It further subsumes the Chain Uniformity Condition and does away with a blatant exception to it in BPS, namely that head-adjunction and Incorporation should be excluded in principle, but are permitted in practice. We maintain that as predicted by the typology, head-adjunction is excluded in syntax, which leaves head substitution as the only available head movement, as is required for e.g. V-raising. We propose an implementation of head substitution based on Grimshaw's (1991) notion of Extended Projections and maintain, contra Chomsky (1995), that head substitution does satisfy the extension condition. The last section of the paper explores the consequence of this proposal for the Mirror effects (Baker 1988, 1998) and argues contra Baker that head-adjunction is therefore not the fundamental explanation behind Mirror effects, but head substitution is.

## 1. The limitations of BPS

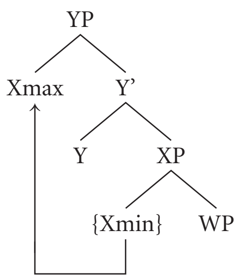
In Bare Phrase Structure (1994) and the Minimalist program (1995), Chomsky proposes to eliminate stipulative aspects of X-bar theory in order to derive core properties of phrase stucture from bare essentials of grammar. In particular, Chomsky seeks to subsume the exact application of Merger and Move operations from primitives of grammar or conceptual necessity, hence obtaining a more explanatory typology of syntactic objects resulting from those operations. The objective is quite stimulating from an explanatory standpoint in that it provides new perspectives for evaluating the overall economy, efficiency, and predictive power of the architecture of the computational system under consideration.

The basic operations involved in forming syntactic objects are Merger and Move. Merger is uncontroversial, it is an asymmetric operation where one of two selected items projects and defines a new constituant. The operation Move, on the other hand, is much more involving and implicates a number of conditions that chains must comply with, as shown in (1).

(1) a. C-command must hold between each link of a chain
 b. Move (chain formation) must obey Last Resort
 c. Chain Uniformity: A chain is uniform with respect to Phrase structure status

While (1a) is uncontroversial, (1b) is a working hypothesis and (1c) a stipulation. Chain Uniformity Condition (1c) in particular plays a very important role. For instance, it excludes the derivation in (2), that is, movement of a maximal head to a specifier, or adjunction to a maximal projection. In both instances, the problem is that the trace of X, shown in curly brackets, is minimal at the tail of the Chain since it projects, but X is maximal at the head of the chain where it does not project. Hence, phrasal Chain Uniformity is violated.

(2)

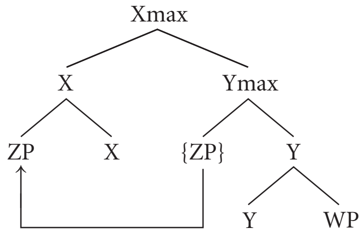



What (1) properly allows however is movement of an XP targeting some YP, by substitution or adjunction and where the target projects. For instance, if the trace of X in (2) were an XP instead of an Xmin, the entire chain would contain XPs and be uniform, provided that YP projects.
Yet, other derivations such as (3) are not directly excluded in BPS.

(3) a. Movement of Xmax to a head Y by adjunction or substitution.
 b. Movement of a non-maximal head Y to a target Xmax and projection of Y.

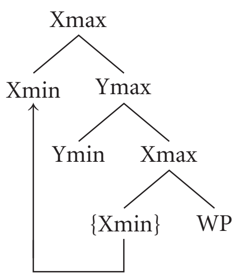
Consider first the case of an Xmax adjoining to a head, as shown in (4). Here a maximal projection ZP adjoins to a head X, an operation generally held as excluded. ZP originates from a specifier in this particular configuration, but it could also be a direct object of X with the same result. This type of operation is not excluded by any condition in (1a–c). In particular, the Chain formed by ZP is perfectly uniform as ZP does not project at the target.

(4)



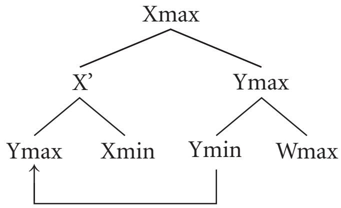

Again, nothing prevents such operation, which is unattested. As for case (3b), shown in (5) below, it is only questionably excluded. In fact, (5) along with example (8) below, represents one of the fundamental generalizations that must be derived about phrase structure and syntactic objects: Only the target of Move can project. In the case of (5), an Xmin targets a Ymax projection, and X projects instead of Ymax. The resulting chain is uniform since X is minimal at both ends of the chain, so the syntactic object is well formed. To prevent such case from occurring, Chomky resorts to condition (1b) Last Resort, stating that movement of Xmin cannot be triggered to check or saturate a property of Y. However, notice that an Xmin could plausibly be moving to assign or verify Case for instance, or to be predicated of Ymax, in which case satisfying Last Resort and yet, these options are clearly not attested. Something more fundamental than Last Resort must prevent a structure like (5) from emerging.

(5)




Another important case is head-to-head adjunction, or Incorporation, as shown in (6).

(6)



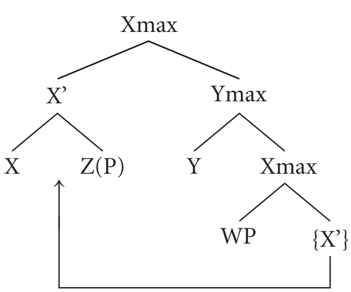

Condition (1c) directly excludes such operation which is abundantly used in the literature.[2] Adjunction of Ymin to Xmin actually makes Ymin maximal at the head of the chain, as Xmin projects. To prevent (1c) from excluding head-to-head adjunction, Chomsky adopts the following ad hoc exception that prevents word internal structure from being subject to conditions of the computational system at LF, (Chomsky 1995: 322).

(7) At LF, X-zero is submitted to word-independent process WI where WI ignores principles of the computational system within X-zero.

This exception is non minimalist in spirit, as it runs against the assumption that the computational system is constant throughout interfaces and for all objects (including heads). As for head adjunction in syntax, Chomsky's assumption that verbs are inserted fully inflected seems to alleviate the problem, though head adjunction is still, in fact, required for Feature Checking Theory, as we discuss in Section 3. Furthermore, Chomsky's (ibid: 334–340) discussion of linear order and Kayne's (1994) LCA suggests that head-to-head adjunction applies in the computational system as well. In sum, head adjunction, or Incorporation, should be excluded in syntax by (1c), but Chomsky proposes (7) to allow it at least at LF, and the issue remains somewhat open in syntax.

A final operation, not excluded by any condition in (1), is shown in (8). The case illustrates an intermediate projection movement where the intermediate projection projects, not the target.

(8)



This derivation respects Chain Uniformity and c-command, and arguably Last Resort (similarly to example (5)). What excludes (8) in BPS is the stipulation that intermediate projections are not visible to the computational system. However, there is no conceptual necessity behind this assumption, and it crucially depends on what is defined as an Xmin or Xmax. Chomsky (1995: 245) defines an Xmin as "...a *terminal element (*selected from the lexicon), with no categorial parts". According to this definition however, the assumption that there is no intermediate projection movement must be ignored when, e.g. V-raising is involved. For instance, if a terminal node Xmin targets and Moves to some terminal Ymin in syntax or LF, it will create a new non-terminal head constituent Y. Yet, this newly created intermediate Y is what is being moved (copied) when multiple head movement applies (see Section 3 for further discussion of the technical details of V-raising under Feature Checking Theory). Thus under Chomsky's own definition of Xmin, there *is* intermediate head projection movement, which implies that somehow the CS must be able to recognize a form of intermediate projection. But again, this significant exception is ignored by the CS under special statement (7). This casts serious doubts on excluding intermediate projection visibility as a general, fundamental property of the CS, hence leaving the exclusion of (8) on questionable grounds. Notice finally that (8) is similar to (5) above insofar as both would involve the projection of the moved element, not the target. That the target always projects is unquestionably a fundamental property of the Move operation and it would be desirable to derive it from a unique source in the CS.

In sum, the typology of syntactic objects in BPS is non-exhaustive in that, on one hand, it doesn't exclude certain unattested objects such as (4), (5) and (8) and, on the other hand, allows for certain syntactic objects such as (6) (head-adjunction) that should be excluded under general assumptions, but are permitted as exceptions under the special condition (7). In fact (7) suspends the application of Chain Uniformity to $X^0$ altogether in the CS, which in turn either questions the status of Chain Uniformity, or the accuracy of an analysis that includes $X^0$ and XP movement under the same Move operation, but treats them differently when it comes to the application of fundamental assumptions such as Chain Uniformity. The typology also relies on the unrelated notions of Chain Uniformity and Last resort to derive the fundamental property that targets always project for both $X^0$ and Xmax movement. The Chain Uniformity Condition itself is basically stipulated.

A further and final remark is that the primitive notions of Xmin and Xmax are not a crucial part of the typology itself, though they play a crucial role w.r.t. Chain Uniformity Condition. Why are there only Xmax and Xmin and how does this distinction bear on the typology of syntactic objects as a whole? I will propose in Section 2 that the very way in which the CS is able to draw a distinction between Xmin and Xmax projections, that is from computing asymmetrical dominance relations between all projections, plays a determining role in defining the entire typology of syntactic objects.

## 2. Asymmetry and Xmax and Xnon-max

Following Di Sciullo (1999), we adopt the view that all structural relations are asymmetric in nature, the Asymmetry Hypothesis.

(9) Asymmetry Hypothesis.
"Any structural relation, be it primitive or derived, is asymmetrical"

Asymmetrical relations are in turn defined as follows:

(10) R is Asymmetrical = (Vx) (Vy) (rxy $\supset$ryx)

Two types of asymmetrical relations can be distinguished, *primitive* or *intrinsic* relations, which are typically structural, and *derived* relations, which are typically grammatical. Hence structural relations of dominance, immediate dominance, sister-contain (but not c-command) and precedence are intrinsically asymmetrical. Other relations, e.g. head-complement and spec-head agreement are asymmetrical by the role that one element plays with respect to the other, e.g. a head selects its complement, a specifier checks features of the head.

Assuming that asymmetry is fundamental in grammar, we explore the nature of the role it might play in defining syntactic objects. Our contention is that primitive asymmetric relations of dominance and immediate dominance are central not only in determining the nature of possible projections, i.e. the basic Xmax and Xnon-max distinction that we propose, but also in restricting the entire set of syntactic objects resulting from the Move operation. The driving idea behind the present proposal is that syntactic objects can only be interpreted by the CS and the LF interface through *primitive* asymmetrical relations, i.e. asymmetry defines and limits the nature of syntactic objects.

Let us first assess which primitive structural relations would be relevant in distinguishing possible types of projections. To answer this, consider first what is a projection. Under minimalist assumptions, projections are sets of formal features which do not have a primitive but only a relational categorial status. Further, projections are endocentric, expanding from lexical items to more complex constructs, in bottom-up fashion. The latter property of projections readily excludes relations of *sister-contain* and *precedence* as relevant in defining categorial distinctions in relational terms. Indeed, given that projections are endocentric and defined in relational terms, only dominating and dominated projections are relevant. That leaves *dominance* and *immediate dominance* as the remaining and relevant primitive asymmetric relations. The question whether both definitions are distinguished by the CS is of interest, but as it turns out, the result of our proposal is the same whether only dominance is considered, or also immediate dominance. Indeed, immediate dominance is but a subcase of dominance and our proposal below only requires dominance. Now, given that projections are essentially bundles of features defined in relational terms, it further follows that two given projections can be either identical or non-identical, feature-wise. We thus obtain the following logical possibilities of distinguishing projections in asymmetric relational terms based on strict feature identity:

(11) a. If Y (immed.) dominates X, and the features of X $\neq$ Y, then X is maximal
b. If Y (immed.) dominates X, and the features of X = Y, then X is non-maximal

The relative notions of *maximal* and *non-maximal* result from computing the identity of a projection w.r.t. dominating and immediately dominating projections. In other words, the Xmax vs Xnon-max (henceforth XNmax) distinction logically derives from the fact that a given projection is defined in relational terms of (immediate) dominance and feature identity.[3]

This proposal is distinct from BPS in several respects. Firstly, notice an XNmax is not equivalent to an Xmin. The notion XNmax does not make a distinction between a terminal Xmin, a non-terminal Xmin and intermediate projections (see example (8) and related discussion on those distinctions). Section 3 develops an analysis of the Mirror effects which supports this conclusion.

Secondly, this proposal is also distinct as it does not determine per se whether a given projection is *overall* an Xmax or an XNmax. This can only be assessed once *each* and *all* dominating projections are considered. Then two scenarios can occur. Under one scenario, the X projection will have an unequivocal Xmax/XNmax status as determined by *each* and *all* dominating projections. Under the other scenario, the X projection would have an ambiguous status, being simultaneously defined as Xmax w.r.t one or more projections, but as an XNmax w.r.t. one or more projections. We come back to examples of such cases directly, but notice first that Chomsky (1995) also considers the possibility that a projection may have a double status, w.r.t. Xmax and Xmin.

Chomsky maintains that nothing prevents such a scenario and speculates that clitics may be an instantiation of such projection, as they seem to have both properties of Xmax and XNmax. We differ fundamentally from Chomsky on that view. We refer to Sportiche (1998) who argues that the alleged double categorial status of clitics is misleading and proposes a unified analysis under which clitics are heads and that the Xmax properties of clitics follow from operator movement.

More specifically, I assume that the computational system and in particular, the LF interface, cannot tolerate ambiguity, be it semantic or relational. That assumption seems motivated on interpretative grounds. Given that phrase structure is still available at the Conceptual/Intentional interface, and given that ambiguities are resolved at that level, each phrasal projection must therefore be unequivocally interpretable as Xmax or XNmax. For the strict purpose of discussion, I will refer to this assumption as the Non Equivocal Requirement, bearing in mind that it arguably needs not be stipulated and follows from the non-equivocal nature of the Conceptual/Intentional interface and the CS.

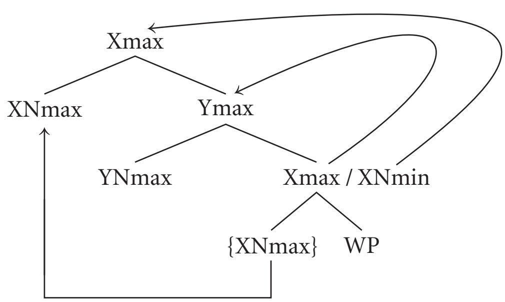
(12) Non-Equivocal Requirement (NER)

A syntactic object must be unequivocally determined as Xmax or Xnmax

Let us now consider direct consequences of our proposal for the typology of syntactic objects.

A first consequence, shown in (13) is that the generalization that the target of Move always projects follows directly and from a unique, primitive property of projections, i.e. the fact that they are defined in relational terms.

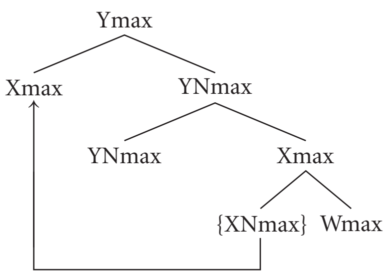
(13)




The derivation in (13) shows the earlier problematic case of an XNmax projecting at the target, hence being well formed w.r.t. Chain Uniformity. Under (12) this derivation is ruled out as one projection is both defined as an XNmax and an Xmax. The culprit in (13) is the projection of X immediatly dominating WP: it is an Xmax by virtue of being immediately dominated by Ymax, but it is also an XNmax by virtue of being dominated by a projection of itself, namely the new Xmax created by projecting XNmax at the target of Move. The phrasal status of that projection is therefore equivocal and uninterpretable, according to (12).

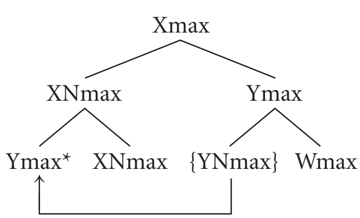
A second consequence of our proposal is that the basic effects of the Uniformity Condition on Chains are also subsumed under (11) and (12). Basically, a Chain is a syntactic object, though a *discontinuous* one. Following Chomsky (1995) for instance, a Chain is a series of copies in c-commanding relations. All and each link is a subpart of a single syntactic object, a Chain. It follows from (12) that a Chain and all of its links must unequivocally be interpretable as either Xmax or XNmax. In other words, chains are subject to the same conditions as Xmax and XNmax. Consequently, the derivation in (14) repeated hereunder for convenience, where XNmax targets Y, Moves and Y projects, will crash as the links of the chain are not of the same type, i.e. the chain is equivocal and uninterpretable. Whereas in BPS, the CUC is stipulated, its basic effects now follow from the general definitions of projections, i.e. Xmax and XNmax, through asymmetrical relations.

(14)




Since our proposal subsumes the effect of Chain Uniformity, it still also excludes head adjunction in the computational system, as shown in (15). The head chain created by adjunction is ambiguous w.r.t. its XNmax or Xmax status.

(15)




Head adjunction also poses a problem for c-command. Chomsky (1995: 334–339) discusses alternative ways of defining c-command so that it holds for head adjunction. Even Chomsky's (1999) notion of *sister-contain* which has the advantage of being a primitive asymmetric relation, still makes adjunction to a head violate locality between chain links.

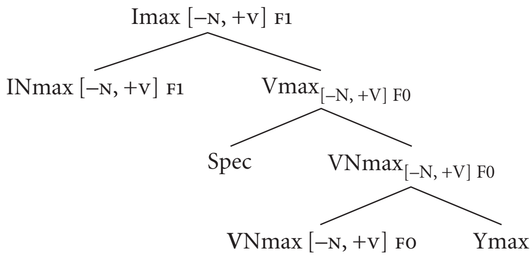
In sum, head adjunction should be excluded by Chain Uniformity, or NER under our proposal, and also by c-command or the sister-contain relation. Keeping with the strongest possible hypothesis, I will henceforth assume that adjunction to a head is excluded in the CS and LF. Of course, this raises a crucial question: How are V-raising and feature checking performed, at least in the CS before spelling out where bare output conditions force pied-pipping of lexical material along with features?

The only other option left for V-raising at this point is substitution. But Chomsky (1995) excludes substitution under the extension condition, which requires the result of Move or Merge to extend a phrase marker. Chomsky maintains that the extension condition is supported on empirical grounds, i.e. provides a version of the strict cycle and prevents raising to object (ibid: 190). These empirical arguments involve XP movement however, and head substitution is discarded on general assumptions. Incidently, Chomsky points out that head adjunction does not extend the target and violates the extension condition, but maintains it on the basis that it does not involve cycle effects or raising to object. Notice incidently that head substitution is not subject to raising to object either, and whichever cycle argument is applied to head substitution could just as well involve head adjunction to an empty head through Merger. This really leaves the question open whether there is real motivation in allowing head adjunction, but excluding head substitution on the basis of the extension condition.

Notwithstanding the latter remarks, a further examination of how head substitution *could* apply suggests in fact that the extension condition *is* satisfied. Since head substitution is not formally discussed in BPS, let us explore an implementation of it.

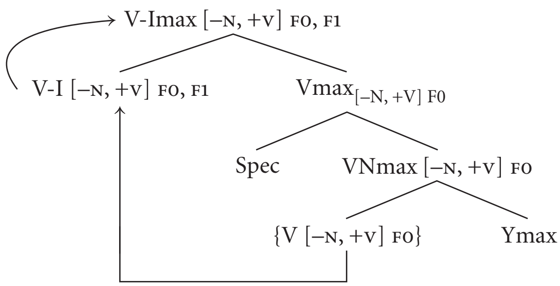
Notice first that head substitution, in contrast to head adjunction, obeys the sister-contain or strict c-command relation between the head of the chain and all its links. Secondly, suppose that it can be maintained that under substitution the head actually projects at the target, then all links of the resulting Chain would be of the same projection level, i.e. the chain would respect NER (Chain Uniformity). And in fact, the idea that a verbal head projects at the target after raising is in essence a logical consequence of Grimshaw's (1991) Extended Projections theory. Grimshaw distinguishes between lexical feature N and V and functional features of level F1, F2, and so on. All categories are specified for lexical and functional category. For instance in (15), V and INFL are both lexically specified as [-N, +V] according to Grimshaw, but V is an F0 functional category while INFL is F1. Hence, apart from the functional feature contribution of a given functional projection, all the functional projections in the extension of a lexical category share the same lexical features.

(16)




If the terminal VNmax were to subsitute under INFL in (16), the result shown in (17) could be conceived as the combination of all features, a merger of the features of V and then of INFL. The newly created category is different from V alone, as it now contains an additional functional feature F1, but it nonetheless remains an extended projection of V by virtue of sharing the same lexical features as V. Such features merger under substitution is arguably possible because both V and INFL share the same lexical features. Yet V and INFL are different projections as they do not share all features, i.e. they differ on at least F1. Hence V does project after raising, as its features combine with the features of Infl to define the new projection.

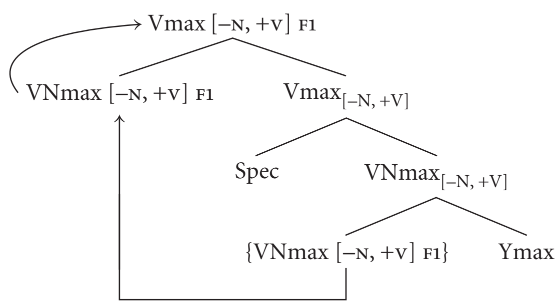
(17)




The consequence is that NER (Chain Uniformity) is preserved by head substitution, as the head of the chain projects at the target and is therefore nonmaximal, as all other links of the chain. In Section 3 and directly below, I modify parts of this analysis so that head substitution obeys the extension condition and provides an account for the Mirror effects. But the core idea remains that the head projects at the target. In sum, head substitution, contrary to head adjunction, meets two fundamental properties of Move: C-command between chain links and the NER (Chain Uniformity).

Let us now return to the issue of the extension condition. Strictly speaking, head substitution still does not comply with the extension condition, even though it projects at the target. And the reason for that is that (16) presupposes the projection of Infl prior to V-raising. But if we consider (16) in strictly derivational terms, it turns out that head substitution is effectively extending the target. In fact, head substitution can be construed as a standard application of Move. Suppose, following Chomsky in that respect, that verbs are inserted in the CS fully inflected, thus as bundles of lexical and functional features. Suppose further that each functional feature on a verbal stem must project its own projection in order to be interpretable and/or checked through spec-head agreement. Then, given that flexional features are extended features of V, we can reanalyze (17) as a case of self-attachment, i.e. VNmax targets its own maximal projection, Move applies and the lexical features of VNmax, along with the functional feature F1, project at the target. This is shown in (18).

(18)

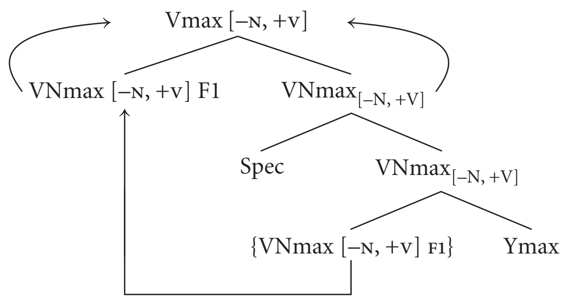



The application of Move now satisfies the extension condition. In fact, there is no substitution in derivational terms (but there is in the representational ones), as there is no head existing prior to Move. But for the mere purpose of differentiating this application of Move from head adjunction, we will still refer to it as head substitution for the remainder of the paper. Notice further that the V head-projects a different category at the target. VmaxF1 is a different projection than Vmax, i.e. these two projections differ in at least feature F1 and are therefore distinct according to (11) above.

This analysis has further interesting consequences. First, notice that the target does not project in (18), but the moved head does. This derivation thus seems to run against the generalization that the target of Move doesn't project. But notice that this generalization does not follow from conceptual necessity and what really matters is to properly exclude the cases where the moved element should not project. As we discussed earlier, (11) properly excludes such cases from a unique source, but we now identify one instance allowed by (11) where the target does not project, and the derivation converges. And coincidently, this option seems to be crucially required by the CS as the only way to perform head movement in accordance with fundamental assumptions about Move.

In fact, there is one more way in which the CS presumably makes use of that option. For instance, suppose once more that self-attachment applies as in (17), but that F1 does not project at the target in this instance. Indeed, nothing forces F1 to project immediately, as long as it does prior to LF interface for interpretability or before spell-out for feature checking. The result would be (19) where either of the target or the moved element could project, as the features involved are exactly the same.

(19)




This scenario allows the option of projecting either the target or the moved head, and the resulting chain is well formed w.r.t. the extension condition, c-command and the NER. Then the question is: Is such a syntactic object attested? Interestingly, this object has all the propeties of a Larsonian VP shell: a verbal head moves to an empty (light) verb position, which under our analysis is actually created by Move. This new head projection could project further and provide a new specifier position. In sum, the option of self-attachment within our proposal leads to an analysis of head substitution which obeys the extension condition as well as predicts the existence of VP shells as well-formed syntactic objects.[4,5]

Summarizing the discussion on head movement, we adopted the firm position that the CS is uniform throughout the CS and interfaces and applies equally to all Move operations. This led us to exclude head adjunction which violates the C-command requirement between chain links, the NER (Chain Uniformity) as well as the extension condition. We developed an analysis of head substitution, based on insights in Grimshaw's theory of Extended Projections, which satisfies all these requirements and is therefore the logical choice for the CS. Section 3 refines this proposal further in developing an analysis of the Mirror effects and features projection.

Summarizing this section, a typology of syntactic objects was developed on the basis of the Asymmetry Hypothesis (Di Sciullo 1999), according to which all relations must be asymmetrical. Syntactic projections being endocentric and defined in relational terms following minimalist assumptions, the primitive asymmetrical relation of (immediate) dominance defined the basic distinction between XNmax and Xmax. The resulting typology has the following properties: It subsumes the effects of the Chain Uniformity Condition; It is exhaustive in that, contrary to BPS, all combinations of adjunction and substitution of Xmax and XNmax are covered; It is exceptionless, in that no operation that would violate a fundamental property of Move or the extension condition is allowed under special conditions. In particular, it excludes head adjunction in the CS and it subsumes the generalization that the target of Move projects, except in one instance predicted by the analysis which corresponds to the only strategy available for V-raising (namely a derivational variant of head substitution) and also provides the source for VP-shells.

Two predictions of this typology are particularly striking: the absence of any head adjunction through Move, and a consequence that we have not touched on yet, namely that intermediate XNmaxs are in principle visible to the computational system. The next section explores one particularly relevant phenomenon linked to these assumptions, the Mirror effects. It will be argued that the Mirror effects are not a result of head-to-head adjunction, as proposed in Baker (1988, 1998), but rather, the result of head substitution involving intermediate XNmax projections.

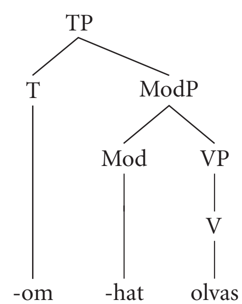
## 3. Head movement and the Mirror Generalization

The cross-linguistic work of Cinque (1999) on the functional structure of the clause reveals a particularly rich make-up and, very importantly, one that doesn't vary very much cross-linguistically. Cinque's work also supports the Mirror Generalization which refers to the fact that inflectional morphemes expressing e.g. tense, mood, and aspect on verbal heads stand in a mirror image to their corresponding functional projections. This Mirror Generalization is commonly held to follow from Baker's (1988) theory of Incorporation, in particular through multiple left adjunction. However, as Cinque points out, and Brody (2000) strongly argues, left-head adjunction does not provide an explanatory account of the generalization. This section of the paper reviews some of Brody and Cinque's arguments and propose an alternative account of the Mirror Generalization that doesn't rely on head adjunction. Our claim is that the Mirror Generalization reflects one of two grammatical strategies to project functional heads into functional projections within the allowed typology of syntactic objects, which, under our assumptions, excludes head-adjunction but allows head substitution.

### 3.1 The Mirror Generalization and head adjunction

The Mirror Generalization refers to the fact that flectional morphemes (tense, aspect, modality, voice, and agreement principally) are in opposite order, or mirror order, with respect to the hierarchy and precedence of their corresponding functional projections in the clause. The following example from Turkish is discussed in Brody (2000), citing Cinque (1999), and shows a typical case of the Mirror Generalization.

(20) Morphological order of Tense, modality and V: o*lvas-hat*-om
read-permissive-1sg present




The Mirror Generalization is often held to follow from Baker's Incorporation theory (1988, 1998), which relies on three fundamental assumptions.

(21) a. Adjunction is always to the left of the targeted head (Baker 1998: 29); "If X and Y are $X^0$ and X is adjoined to Y in the syntax, then X precedes Y in the linear order".
b. Head-movement is subject to the Head movement Constraint.
c. Head adjunction excludes Excorporation.

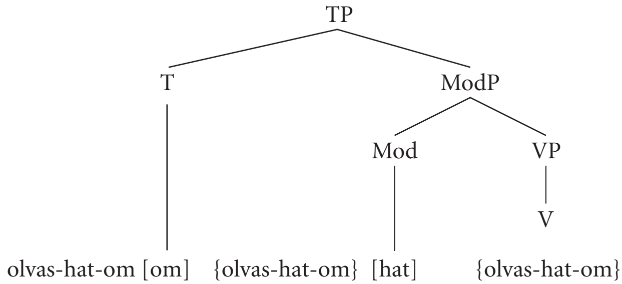
According to these assumptions, the mirror order follows from successive head adjunction, where the verb raises and left-adjoins to the next head up, and then this new verb coumpound raises and left-adjoins to the next head up.

However, author such as Di Sciullo and Williams (1987a, b Section 3.6), Speas (1991), Cinque (1999), and Brody (2000) among others, discuss compelling arguments showing that this account is inconclusive. In particular, two of the basic conceptual assumptions in (20) are questionable. Firstly, Brody points out that the exclusion of excorporation for head movement, which is crucial to account for the strict reverse order and to prevent the form **olvas-om-hat*, is in direct contrast with adjunction to XPs, which crucially allows excorporation. There is indeed no independent evidence that supports such special status for head movement, and Koopman (1994) argues that excorporation is possible. Secondly, notice that restricting adjunction to the left is also dubious. Even Baker's (1998: 33) claim about Sora, a Mundu language of the Indies showing inverted mirror effects, suggests that left adjunction is not universal: "These facts are elegantly explained if the direction of adjunction performed by head movement is simply reversed in Sora". So according to Baker, left head adjunction is not a constant assumption either.

In addition to these conceptual questions (see Brody (1998) for other problematic issues pertaining to head adjunction w.r.t. c-command and relativized minimality), Brody also points out serious shortcomings for feature-checking theory. Let us briefly consider some of these problems and add some comments to them, as they are directly related to the alternative proposal developed in the next section.

A first issue resides in the fact that not all verbal forms showing a mirror order actually move in syntax, as amply argued on the basis of French and English (Emonds 1978; Pollock 1989; Chomsky 1995). Verbs would only move to INFL overtly in French, yet both languages display a mirror effect with respect to flectional morphemes. If the mirror order results from successive head adjunction in syntax, English then remains a mystery. Chomsky's (1995) feature checking theory provided a solution to this paradox. Basically, verbs in both languages are inserted fully inflected in situ and Move makes copies of the inflected verb and adjoins it in each head position of the head chain. Each copy allows feature checking, but only one copy surfaces at spell out (i.e. under VP in English, and under INFL in French). The final result for e.g. Turkish is the following (from Brody 2000: 36), where items in square brackets indicate non-spelled out heads and those in curly brackets, non-spelled out copy-traces.

(22)

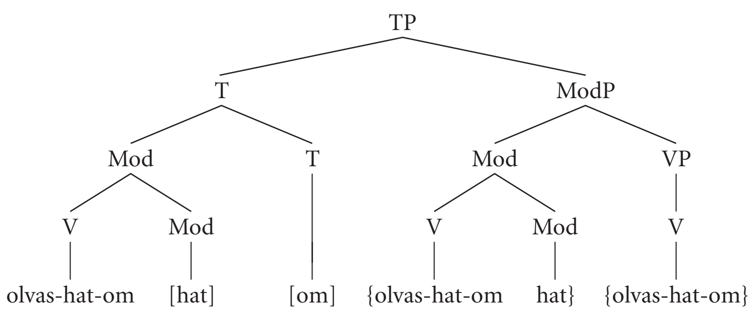



This derivation however, does not warrant that the Mirror Generalization follows from left adjunction. As Brody points out, new stipulations are required to explain the mirror effect, namely:

(23) a. Feature checking must start with the most embedded suffix.
b. Feature checking must follow the order of the suffix, from the most to the least embedded.

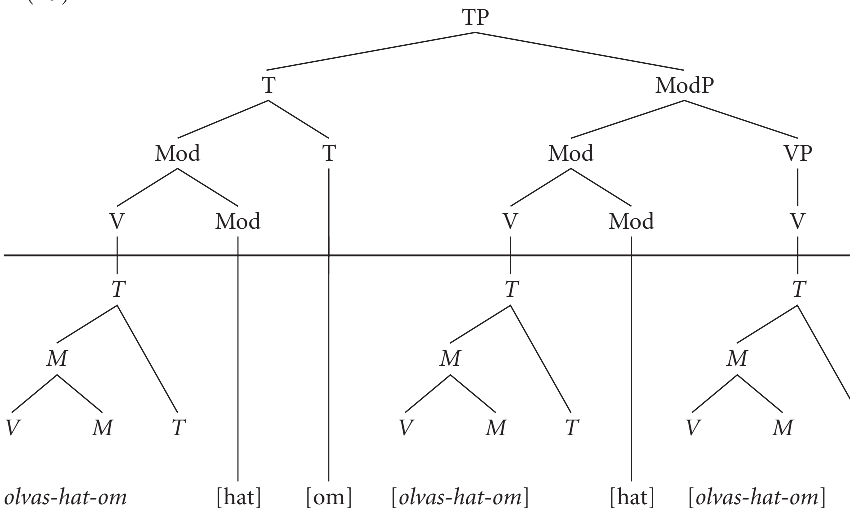
Independently of the question whether these stipulations are a satisfactory account of the Mirror Generalization, the fact is that left-adjunction has no bearing anymore on the account of the generalization itself. There are further problems as well. Brody rightly points out that the complete structure for (21) is actually more complex. General assumptions about words, phrase structure and chains require that words and chain members be constituent nodes, that nodes be labeled by one of their constituents, and that only tails of chains may be labeled (i.e. only the target of Move projects). The combination of these assumptions requires that entire copies of traces form a constituent with the head of each functional category, itself not spelled out. The end result is the following structure from Brody (ibid: 36).

(24)




To this syntactic representation, we must further factor in the word-internal structure of the word, which in essence duplicates the syntactic structure. Assuming a morphological right adjunction word formation, the final representation is the following (with the word-internal structure italicized and separated from the syntactic structure by a horizontal line, for more transparency).

(25)




Let alone the amount of structural and feature specification redundancy in this representation that Brody discusses in details,[6] plus the fact that left adjunction still has no bearing on the mirror order, there are at least two additional oddities worth pointing out. Firstly, notice that the uppermost label of the inflected word, *T*, is never inserted under a node of the same type, but under a V node. Secondly, none of the features of the word is actually checked in a local fashion with a functional head, e.g. under sisterhood or immediate dominance or "spec-head" agreement. In the case of *T* for instance, its feature must be checked through the V and Mod syntactic nodes before actually reaching the syntactic T node. Notice that this structural distance would only increase as the number of functional morphemes would, i.e. each new head would imply more structural distance between affixes and functional heads.

A similar remark can be made about the morpheme *V* and the syntactic head V, however in this case, a solution along the lines of the notion of Relativized Head (Di Sciullo & Williams 1987) is available. Hence the categorial features of *V* and *M* percolate to *T* and are thus available from *T*. This percolation is possible since *V*, *M*, and *T* specify distinct types of features, according to Di Sciullo & Williams. After percolation a stronger locality between features and functional heads would be warranted, but ad hoc assumptions along the lines of (23) would still be required to explain the mirror order of morphemes.

In sum, a closer look at feature checking indicates that left-adjunction has no bearing in explaining the mirror order of suffixes. The notion of Relativized Head says nothing about the internal order of morphemes and features, so the Mirror Generalization can then only follow from other stipulative assumptions such as (22). Further, the notion of Relativized Head, which is of the domain of word-internal features, says nothing about the non locality that persists for feature checking between *T* and *M* and their corresponding syntactic, functional heads M and T.

Summarizing this section, several arguments were presented, in particular from Brody (2000), showing that an explanatory account of the Mirror Generalization cannot be held to follow from successive left head adjunctions, even considering the framework of feature checking theory (Chomsky 1995). The next section develops an alternative account of the Mirror Generalization within the typology of syntactic objects developed in Section 2.

### 3.2 The Mirror Generalization and head substitution

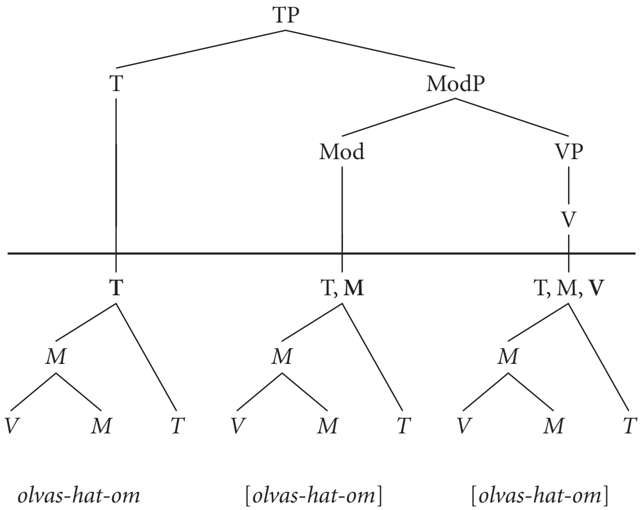
The typology of syntactic objects developed in Section 2 excludes head adjunction in syntax through Move. That considerably restricts the CS in performing Verb-raising under head movement, basically leaving head substitution as the only alternative. In the remainder of this section, an analysis of Mirror effects based on head substitution is explored. In the process, some aspects of head substitution as proposed in Section 2 will be refined, while coming to the following conclusion: Mirror effects reflect the only alternative available to the CS to ensure that each relevant flectional feature associated to a bound morpheme projects its own syntactic projection.

Before developing an analysis, let us review some assumptions about functional features projection. As mentioned at the outset of Section 3, one of Cinque's important conclusion about the hierarchy of clausal structure, in particular for tense, mood, voice, and aspect features, is that there seems to be no cross-linguistic variation. In the Principles and Parameters approach of the Minimalist program, this suggests that the relative order of those features is not ultimately determined in the lexicon, which is the source of parametric variation, but follows from invariant interpretative principles related to bare output conditions of the conceptual/intentional (C/I) interface. For instance, hierachical structure determines the relative scope of functional heads and the fixed cross-linguistic order could result from a universal relative scope interpretation required of those specific morphemes. It is beyond the scope of this paper to substantiate that specific hypothesis, but the fact that there is no cross-linguistic variation is consistent with the assumption that the sentential order of those functional features is not subject to lexical variation but is rather determined by bare output conditions of the C/I interface. Of course, whether a given feature is realized as a bound or free morpheme, or as a suffix or prefix, still is a matter of morphological variation and must be specified in the lexicon. But assuming that Cinque's generalization is right, let us suppose that the relative sentential order of the projections corresponding to those functional features, contrary to perhaps other features such as negation or agreement, is not determined in the lexicon, which would lead us to expect parametric variation, but by bare output conditions at the C/I interface.

The next logical question to ask is what are the options available to the CS to instantiate that relative functional feature order. Again there are few options. Assuming that functional features must reach a relative order for scopal interpretation, this implies that the CS must ensure that every relevant functional feature projects its own scope at the C/I interface. Since scope is determined under c-command, each functional feature must head its own syntactic projection to establish its scope. A straightforward way to instantiate this would simply consist into selecting from the lexicon a morpheme corresponding to a feature and project it directly, through Merger, as a syntactic head. Depending on whether the morpheme is phonologically bound or free, it could or not create a phonological word with adjacent morphemes without changing their relative order, e.g. through phonological cliticization. This is exactly what Cinque suggests is happening in languages of the Bantu family, where "...prefixes appear to directly reflect the order of functional morphemes, with agreement to the left of tense, which is itself to the left of Aspect." (Cinque 1999: 68). Cinque presents evidence to the effect that those prefixes are actually free morphemes and points out that these languages are a challenge to the Mirror Principle which doesn't seem to apply. But in our view, there is no reason for the Mirror Generalization to manifest itself in this circumstance if the mirror order is but a reflection of a movement strategy to instantiate feature projection. The morphemes in the Bantu family are inserted directly as syntactic projections, hence directly establishing their scope, and there is no need to resort to any further strategy that would result e.g. in a mirror order.

Let us now consider a different strategy from direct projection of features into syntactic projections. Another option that the CS must deal with consists into forming complex, fully inflected verbs in the lexicon, created by word formation (presumably head adjunction) and insert them as such in syntax. This is Chomsky's (1995) position.[7] The CS must ensure that each feature of that inflected verb is projected for interpretability. The analysis of V-raising proposed in Section 2 provides a proper way of executing this. Consider a first attempt at a derivation, where specifiers positions have been omitted as we focus on head movement.

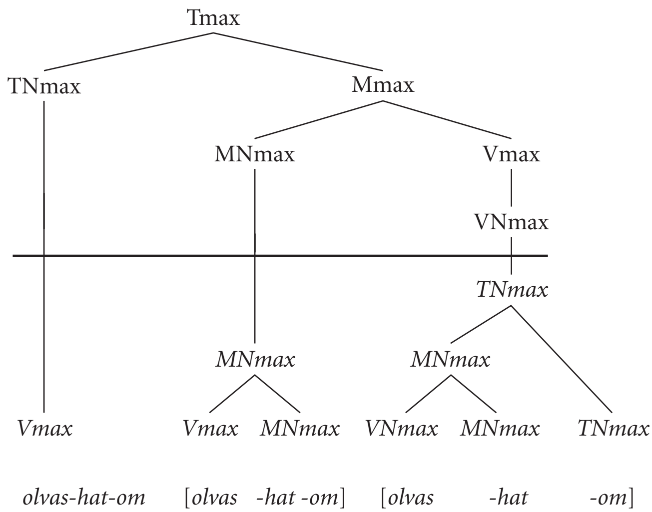
(26)




Head-substitution of the inflected verb was performed through copying. Assuming percolation of features through relative headedness, each feature of the verbal complex is available at the topmost node of the word and can project directly from there into its own syntactic projection. Though a positive result is achieved, nothing however warrants that the order of suffixes should be in a mirror order. This is indeed the conclusion we reached in the last section when we discussed the use of relative headedness or feature percolation.

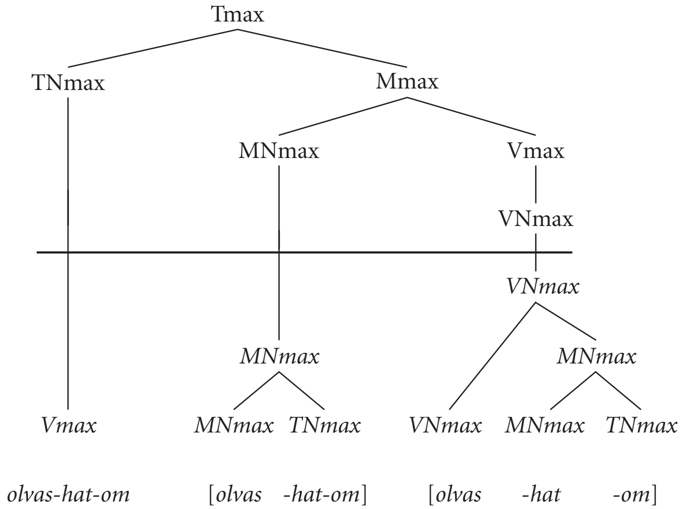
Let us therefore explore a different option offered by the typology in Section 2, one that will not use relative headedness. Recall that the typology only makes a distinction between Xmax and XNmax, i.e. does not make a stipulated distinction between Xmins that are lexical items and Xmins created by head movement. XNmax is simply any projection visible in syntax but which is not maximal, as determined by asymmetric dominance relations. Suppose therefore that the CS can identify flectional morphemes as XNmax, which seems reasonable for those flectional morphemes that have syntactic projections. Assuming that individual flectional morphemes are XNmaxs, they can undergo Move (head substitution) as long as the constituent structure of the word is respected. In that perspective, consider the following derivation.

(27)



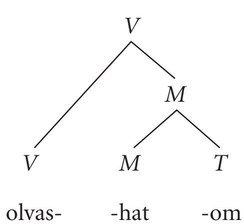

This derivation has proceeded through successive head movement (copy) of a subpart of the word flectional structure. Notice that even though only a subpart of the word structure moves, the whole phonological word is pied-piped for PF convergence. However, this derivation arguably crashes for a number of reasons. For one, the head of Tmax is a *Vmax*, i.e. the feature *T* is not projecting. In fact, the status of the lowest *VNmax* is similarly questionable. Since we are allowing access to word internal structure, relative headedness should presumably not be used and consequently, the feature of *VNmax* cannot percolate to the highest node *TNmax*. Thirdly, notice also that the projection status of V changes in the copying process of *MNmax*, passing from *VNmax* to *Vmax*. Even though *VNmax* is not the head of the chain, that clearly seems an undesirable consequence.

So derivation (26) is excluded on several grounds and, in any case, it still did not provide any explanation for the mirror order. However, the interest of (26) lies in suggesting another option which would provide the results we are seeking. Instead of assuming a right branching adjunction structure of the word, let us suppose an inverted structure, preserving the same morpheme order, and then perform head substitution.

(28)



The derivation again proceeds in copying subparts of the flectional word structure, but this time it converges. For one, given the constituent structure posited, each feature now directly projects into its own syntactic projection. In addition, all chains respect the c-command requirement as well as the NER (Chain Uniformity), and no projection level change results from a copying process. Notice further the dramatic decrease in computational complexity and feature redundancy of (28) if we compare it to the head adjunction analysis above in (25). But the most striking result is that the mirror order of morphemes now follows in a principled way. Indeed, the only way each feature can be directly projected, while preserving the constituent structure and the fact that those morphemes are suffixes, is if the morphemes are in the mirror order from the functional projections. Any other order of the morphemes would necessarily lead to at least one feature not being projected under its corresponding functional projection.

This analysis provides the expected results, but it relies heavily on the assumption that the word structure of inflected morphemes can be inverted, as shown below.

(29)


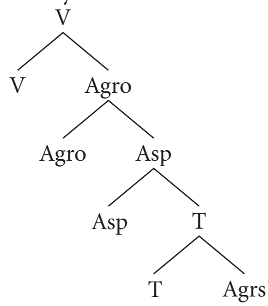
This morpheme structure does not reflect the hierarchical structure of the corresponding sentential projections, and it would seem hard to justify that the verb selects Mood which in turns selects Tense. And indeed, it is not our contention to do so. Recall from the outset of this section that one logical consequence of Cinque's findings is that the sentential order of certain flexional heads -those involved in the mirror effects and which are not subject to cross-linguistic variation- is not determined in the lexicon but by bare output conditions at the C/I interface. In other words, apart from specifying that in Turkish, those morphemes are suffixes and that they are extended features of verbs in the sense of Grimshaw, we can assume that the lexical component has nothing more specific to say about the relative or hierarchical order of those suffixes in the word. All that is required is that the way those morphemes will be adjoined to one another will ensure that the CS will be able to project them so that they meet the relative order required by bare output conditions at the C/I interface.

The analysis above extends in a direct way to cases of Athakaspan languages such as Navajo. As argued in Speas (1991), Navajo and other Athapaskan languages display inflectional morphemes as prefixes. What is striking however is that those prefixal functional morphemes show the same relative order found in other languages such as Greek, Finnish, Basque, and French. The only and crucial difference is in the position of the verb stem.

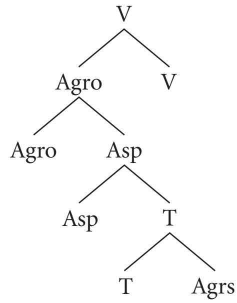
| | | | |
|---|---|---|---|
| (30) | a. | Greek, Finnish, Basque: | verb stem-OAGR-ASP-T-SAGR |
| | b. | Navajo & Athapaskan: | OAGR-ASP-T-SAGR-verb stem |

Notice first that the mirror order follows directly under our analysis, despite the morphemes being prefixes. In fact, there is no significant difference in the derivation in Navajo and Greek or Basque, the only difference being that the bundle of functional morphemes is in a different position w.r.t. the verbal stem, a morphological parameter according to Speas (1991). Recall that the relative order and the hierarchical structure of inflectional morphemes among themselves have no bearing on the output of the word, but are only relevant in ensuring a converging derivation at the C/I interface. Hence, the mirror order is still necessary in Navajo, despite the fact that morphemes are prefixes. The structure of flexional morphemes for each language group will be as follows.

(31) a. Greek, Finnish and Basque family

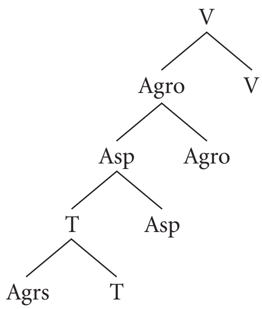



b. Navajo and Athapaskan




Let us now come back to the question of Why there is still a mirror image with prefixes. In other words, Why doesn't Navajo display a non-mirrored order with the following hierarchical structure?

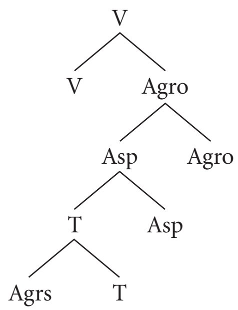
(32)




This word structure would allow unfolding the morphemes under head-substitution and processing with feature checking under strict locality. To our knowledge, no language displays this order with the prefixes being bound morphemes, and not free morphemes as in Bantu languages. Provided that this is a true empirical generalization, we expect it to follow from some fundamental property of the computational system. One interesting possibility is to capitalize on a version of Kayne's (1994) Linear Correspondence Axiom (LCA), for instance as proposed in Chomsky (1995: 334–340). In (30), each morpheme that precedes another morpheme is containing (including) and sister-containing (asymmetrically c-commanding) that morpheme. In contrast in (30), there is no sister-contain relation that corresponds to linear precedence, i.e. the LCA is violated. This again leaves (31) as the only possible hierachical orders of functional morphemes that will ensure a converging derivation. The LCA therefore provides an explanation for the presence of a mirror effect in languages with functional prefixes, such as Navajo. Since these functional morphemes are visible to the computational system under our analysis, it is unsurprising that the LCA should constrain their hierarchical structure.

This account in terms of LCA could extend to another potential case scenario. Let us suppose a language such as Basque or Greek, but where the functional morphemes would *not* show a mirror effect. To our knowledge, such language does not exist. Yet under our analysis, there is indeed one hierarchical structure that could exactly make that prediction, i.e. a language with flectional suffixes but with no mirror effects. The word structure would be as follows.

(33)




Unfolding the morphemes through head substitution would ensure projection of each morpheme without requiring a mirror order. However, notice that the flectional morpheme structure is essentially similar to (31) and violates the LCA for the same reasons. So the LCA also excludes this unattested case.

In sum, the discussion above suggests that the LCA is operative in ordering the structure of flectional morphemes, which comes as no surprise under our analysis which maintains that certain flectional morphemes are visible to the CS as XNmaxs. The effect of LCA is to further constrain the possible hierachical structure of flectional morphemes, leaving basically structures such as (31) available for head-substitution, which incidently, necessarily imply a mirror order.

## Conclusion

This paper developed a typology of syntactic objects based on primitive asymmetric relations of dominance, following the Asymmetry Hypothesis of Di Sciullo (1999) according to which all structural relations are asymmetric in nature and asymmetry is a fundemantal building block of Grammar. The typology is a criticism of and an alternative to Chomsky (1994) and (1995). The typology that results has the following properties. It isolates two basic syntactic objects, Xmax and XNmax, which result from assessing the status of each feature projection in relational terms through the primitive asymmetric relation of (immediate) dominance. The typology derives the fundamental generalization that the target of Move projects, except in exactly one case which we argue to give rise to VP shells as proper syntactic objects. The typology also subsumes all effects of Chomsky's stipulated Chain Uniformity Condition. One of the most important consequences of the typology is that head adjunction under Move is excluded in the computational system, which leaves head substitution as the sole head movement operation available. We developed a strict derivational approach to head substitution which makes it compatible to the extension condition. Section 3 addressed the Mirror Generalization and proposes that head-substitution, and not head-adjunction, is the source of the generalization.

The typology developed in this paper raises many empirical issues and puts into question a number of standard assumptions in the minimalist program which only further inquiry will allow to further evaluate. For instance, while the seminal work of Cinque is significant, more research is required to verify the cross-lingusitic validity of the Mirror Generalization and confirm exactly what features are subject to it and why exactly there is no cross-linguistic variation in the order of a subset of fonctional morphemes. Also, the exclusion of head-adjunction in the CS certainly has tremendous consequences for feature checking theory, especially covert Feature raising, which we have not considered. In addition, allowing the CS to have access to part of the word structure runs against the strict lexicalist hypothesis (Lapointe 1980; Williams 1981; Di Sciullo & Williams 1987, and Chomsky 1995), and opens a debate that we have not taken on here due to the limited scope of the paper. However, our goal was to lay out a typology of syntactic objects that takes root in the fundamental building block that we think asymmetry is in grammar, and explore the predictions of the typology without resorting to exceptional conditions and keeping the CS constant throughout, even if this required puting into question standard assumptions.

## Notes

1. This paper has benefited at various stages of development from comments of members of the Asymmetry project, in particular Rosemarie Déchaine, Anna Maria Di Sciullo, Manuel Espanol-Echevarria, Philippe Gabrini, Mohamed Guerssel, Yves Roberge et Mireille Tremblay, from the audience at the Asymmetry Conference and CLA, also from Peggy Speas,

Tom Roeper and anonymous reviewers. This paper was also produced under the financial support of the SSHRC and SFU's Vice-President office.

**2.** Condition (1a) is also excluding head-to-head adjunction, unless adjunction is taken to form segments of category transparent to c-command. This requires special assumptions about c-command, which are far from conclusive (see Brody 2000). See Section 2 for further discussion on this issue.

**3.** Speas (1990:44) proposes as partly similar approach with the following definitions for Xmax and Xmin:

i. X= Xmax iff all G which dominate X are distinct from X

ii. X= $X^o$ iff X immediately dominates a word.

The definition for Xmax is based on dominating projections, similarly to (ii). However, (ii) does not stipulate the overall status of a projection, but merely how it is determined w.r.t. one given dominating projection. See directly in the main text the discussion on NER on this issue. Notice further that the definition for Xmin relies on the primitive notion "word", which does not follow from (immediate) dominance relations and feature identity. $X^o$ is thus similar to Chomsky's Xmin, but not our XNmax.

**4.** P. Speas (personal communication) suggested that this way of deriving VP shells provided an explanation as to why a lexical or functional head can project at most three argument positions within its projection. Hale and Kayser (1999) proposed that a predicate may project at most 3 arguments because there are three distinct argument relations, namely object of, specifier of and predication. However, if the light verb in a VP shell is merely another distinct v-position, the question is why couldn't there be more than one distinct light verb and therefore, more than 3 distinct argument positions. But if VP shells are derived by self attachment as we suggest, the light verb is actually a copy and not a distinct head per se. Assuming arguments of a given head must be realized in a different structural fashion for purpose of interpretative distinction, the idea that VP shells are the product of self attachment would indeed limit the number of possible distinct structural relations available to three, namely: sister of a head, spec of a head and sister of a projection of a head, i.e. the new specifier created by self-attachement. In other words, nothing would prevent self-attachment to apply more than once, but it could not create more than one new structural relation to realize a distinct argument.

**5.** This analysis of self-attachment also has potential consequences for the existence of multiple spec agreement and the same-feature restriction displayed by the phenomenon. If multiple specifiers are strictly created by self-attachment of the same head, and assuming as in the previous footnote that an asymmetrical relation can only express one formal property within a projection, a spec-head relation could apply more than once in a projection, but only to express the same feature/argument relation.

**6.** Brody points out that there is quadruple redundancy of feature specification that results from specifically applying the assumptions of feature checking theory and left head adjunction.

**7.** Notice that under the typology proposed in Section 2, we can provide a principled explanation as to why that may be the only option: Head adjunction is excluded in syntax, therefore if a language has morphologically bound functional morphemes, the latter will necessarily have to be merged in the lexicon to avoid head adjunction in syntax. The prediction is that languages with bound inflectional morphemes will insert fully inflected forms from the lexicon.

## References


Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing.* Chicago: University Chicago Press.

Baker, M. (1998). *The Polysynthesis Parameter*. New-York: Oxford University Press.

Brody, M. (1998). Projection and phrase structure. *Linguistic Inquiry, 29* (3), 367–398.

Brody, M. (2000). A mirror theory. *LI, 31* (1), 29–56.

Chomsky, N. (1994). Bare phrase structure. *MIT Papers Working in Linguistics, 5*, Cambridge, MA: MIT Press.

Chomsky, N. (1999). Derivation by phase. *MIT Occasional Papers in Linguistics, 19*, 1–43. Cambridge, MA: MIT Press.

Cinque, G. (1999). *Adverbs and Functional Heads: A Crosslinguistic Perspective*. Oxford: University Press.

Di Sciullo, A. M. (1999). The local asymmetry connection. *MITWPL, 35*, 25–47.

Di Sciullo, A. M. & Williams, E. (1987a). Noun incorporation vs. cliticisation. Paper presented at NELS. Amherst: GLSA.

Di Sciullo, A. M. & Willliams, E. (1987b). *On the Definition of Word.* Cambridge, MA: MIT Press.

Emonds, J. (1978). The verbal complex V'-V in French. *Linguistic Inquiry, 9* (2), 151–175.

Hale, K. & Keyser, S. (1999). Bound features, merge and transitivity alternations. In Pylkkanen, van Hout and Harley (Eds.), *Papers from the UPenn/MIT Roundtable on the Lexicon MITWPL, 35* (pp. 49–72).

Koopman, H. (1994). Licensing heads. In D. Lightfoot and N. Hornstein (Eds.), *Verb Movement* (pp. 126–295). Cambridge: Cambridge University Press.

Lapointe, S. (1980). A theory of grammatical agreement. Ph.D. dissertation. U. of Massachusetts, Amherst.

Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry, 20* (3), 365–424.

Speas, P. (1990). *Phrase Structure in Natural Languages.* Dordrecht: Kluwer Academic Publishers.

Speas, P. (1991). Functional heads and the mirror principle. *Lingua, 84*, 181–224.

Sportiche, D. (1998). *Partitions and Atoms of Clause Structure: Subjects, Agreement, Case and Clitics*. London and New York: Routledge.

Williams, E. (1981). On the notions «lexically related» and «head of a word». *Linguistic-Inquiry, 12* (2), 245–274.

# Synthetic/analytic asymmetries in voice and temporal patterns*

Abdelkader Fassi Fehri
Mohammed V University

Among the issues that any linguistic theory has to deal with are the principles and limitations on the asymmetrical mappings between sound and meaning in terms of words and phrases, and typically how grammatical functions (or functional categories; GFs henceforth) arising in lexical or syntactic derivations are mapped into morphonological segments or sequences. Lexeme-based morphology (cf. Beard (1995) was primarily designed to account for a number of these asymmetrical mappings through the Separation Hypothesis.[1] Distributed Morphology (cf. Halle & Marantz (1993)) has refined the picture to include even more complex relationships between lexico-syntactic structure, morphology, and actual vocabulary.[2] But many questions remain to be solved. For example, even though the traditional classification of languages into *isolating, agglutinative,* or *fusional* is presumably not directed to focalize on the existence of purely discrete types, the relevance of these classifying properties has not received notable attention.[3] Thus the one-to-one mapping found in isolating or agglutinative Ls, as opposed to the many-to-one mapping found in fusional (or cumulative) Ls, is commonly treated as if it were of no empirical (or theoretical) significance. If a (relatively) isolating-agglutinating-fusional continuum is true of most (if not all) languages, then the question that arises is: which natural principles and constraints operate in the organization of the continuum?

Best studied traditional mapping relationships have been usually confined to the domain of morphological words (be they roots, stems, simple or complex words), with a one-to-many mappings (when a single function is associated with multiple discrete expressions, or "synonyms"), or with a many-to-one mapping (when multiple discrete functions share only a single expression, which is "polysemous" or "homonymous"). But clearly, the mapping contrasts are not to be found only within word structure, but do distribute simultaneously and naturally on phrase structure, with Ls varying as to whether they use multiple (discrete) words to express GF complexity, or they use only a single word for such expression. I will call the first type of expression *analytic*, and the second, *synthetic*.[4] For example, the various components of a Mood-Tense-Aspect-Voice-Verb chain can be grouped around a single thematic verb (properly inflected), or they may distribute also on one (or more than one) extra auxiliary verb. It is the characterization of this kind of variation that the article will be dealing with.

Most (if not all) languages resort to various mixtures of the synthetic and analytic strategies, but the choices are not free. Evidence will be provided, in fact, for the following claims:

a. the options of analyticity or syntheticity are *disjunctive*, in the sense that no language makes use of both strategies to express the same GFs (and/or meanings). The following disjunctive "global" asymmetric constraint appears to limit the set of possible morphological or formal objects available in any language; call it *Formal Disjunctivity*:[5]

   (1) *Formal Disjunctivity*
   For any formal relation *R*, established between *x* and *y* in a language *L*, where x and y are morphemes or words that express GFs, R in L is *either* synthetic or analytic.

b. the choice available in (1) is *not L-free*, in the sense that it is correlated with other morphologically and/or syntactically detectable properties of L.
c. the *level of analysis* (or word split) in L complies with the hierarchical organization of GFs (assumed to be a UG property; cf. e.g. Cinque (1999)), and it operates from bottom to top, so that when a split occurs at the level of $GF_i$, all GFs that $GF_i$ dominates occur in the word where the latter is found.
d. *Formal Complexity* (= FC), based on Temp(oral) and Agr(eement) complexity essentially, plays various roles in creating verbal analysis in the same L, or across Ls. Analysis occurs only in Ls that exhibit an AgrS/AgrO split.
e. *Voice* belongs to the class of Temp GFs, which can be dominated by a Temp (verbal) projection, or a (nominal) Agr projection.
f. *Nominal* Agr may be used in Ls to form complex tenses, and the nominal or temporal content of verbal auxiliaries provides variation in the analytic/synthetic continuum.[6]

g. choices made by L become *peculiarities* of L, in the sense that they create limitations or exclusivities on what can be expressed in L through grammatical combinations.

In compliance with (1), the article will investigate limitations and constraints on analysis or synthesis through examining the patterns of variation in the expression of Voice and Temporality found in a number of Ls, compared typically with Arabic. It is shown that analysis finds its source in the specific degrees of FC, arising typically in the Temp and Agr systems, and leading to appropriate Splits (in Temp or Agr chains). Voice may also be playing a role in increasing FC (and hence triggering analysis), depending on whether it belongs to finite or non-finite chains. In Section 1, analytic/synthetic patterns of expression of Perfect Tense and Passive Voice in a number of Ls are examined and constrasted with their Arabic counterparts. It is then argued that the contrasts in the Perfect expression can be attributed to an *Agr Split* (into Agr1 and Agr2), which is or is not available for L. The latter variation depends on the *nominal* nature of the auxiliary (in the analytic type), or its *verbo-temporal* nature (in the synthetic type). On the other hand, analytic complex tenses are shown to originate from a *Temp Split,* due to Temp complexity (interacting with Agr complexity). The same is shown to be true of Voice expression, once it is admitted that Pass(ive) belongs to the Temp chain, and it can be either (immediately) dominated by a nominal Agr (i.e. a non-finite participle Agr2), or by a verbo-temporal projection (with no availability of Agr2, and no split in the Temp chain). In Section 2, more precise content is given to the sources of FC, which triggers analysis. Some peculiarities of Ls are then addressed, including those of analytic Pass Perf in Classical Latin, which provide support for the view that Voice increases Temp Complexity. Likewise, the analytic character of complex tenses in Arabic (as opposed to synthesis in Passive and Present Perfect) is contrasted to that of the synthetic Ancient Greek, by taking into account the peculiarities of Arabic verbal inflection. The latter constantly bears Pers(on) finite morphology, and it does not instantiate Agr2. The section also addresses the proposal of *deflexion* as a direct source of analysis (through reanalysis), and rejects it as a potential explanation. The *Splitting analytic theory* is defended instead. In Section 3, further peculiarities of Ls are investigated, including e.g. the existence of imperfective Pass in Arabic, and its absence in French. The latter are related to other peculiarities in the Temp systems of the two Ls. The clear-cut division in Arabic between verbal and adjectival Voice expressions is also examined and compared to the ambiguous categorization of English Pass participles. It is proposed that various attachment Heights of the same form can be used to express multiple GFs, without resorting to multiple morphologies.

## 1. Analysis, Voice, and Temporality

### 1.1 The problem

Consider the following verbal Voice contrasts between Arabic, English, and French, respectively:

(2) qutil-a r-rajul-u
killed.pass-P the-man-nom
a. The man was killed.
b. The man has been killed.

(3) a. The man was killed.
b. The man has been killed.

(4) L'homme a été tué.

In (2), there is only a single verb, which expresses Pass (ambiguously), either in the Past or the Present Perfect tenses (P in the gloss marks Past of Perfect). Internal vowels *u-i* mark the Pass, and alternate apophonically with the active (marked as *a-a* in (5) below, for example). The suffix on the verb marks Past or Perfect, as well as Pers. Thus Voice and Tense (and/or Aspect) are fused in a single synthetic verbal form (although in distinct positions). This form of expression clearly differs from that found in English and French. In English (3), two distinct passive constructions are needed for each Tense, and both of them are analytic. In French (4), the passive verbal expression is also analytic, and it makes use of two auxiliaries, although this form of *passé composé* translates ambiguously either the English Past (3a) or the Perfect (3b). There are no other means to express these tense and voice combinations in the three languages, hence providing an instantiation of the disjunction stated in (1).

Note that a similar contrast is found with active present Perfect:[7]

(5) kataba r-risaalat-a l-ʔaan-a (* ġad-an)
(has) written the-letter-acc now (* tomorrow)
He has written the letter now.

(6) John has written the letter now (* yesterday).

(7) Jean a mangé la pomme en ce moment (* hier).

Observe again that the active Present Perfect in Arabic is synthetic, while its expression requires an auxiliary in both English and French. Why is then such a difference? Why can't Arabic make use of a verbal auxiliary in the Present Perfect, and why is the Passive expression confined within the finite verb (like active Voice is), in contrast to English and French, which must resort to periphrastic verbality, in which the auxiliary carries Tense and Agr (finite) features, and the thematic verb Voice features?

### 1.2 Nominal auxiliaries

The absence of the auxiliary with the Arabic Present Perfect recalls the absence of the Present verbal copula in so-called nominal sentences. Note that in English and French, its presence is obligatory:

(8) r-rajul-u (* y-akuun-u) fii d-daar-i
The man (is) in the house.

(9) The man * (is) in the house

(10) L' homme * (est) à la maison.

These sentences have exactly the same (Present tense) interpretation, although the first one must be verbless, while the other two must contain a copular verb. Since it is normally agreed that English and French have no overt temporal morphology in the Present, it is reasonable to think that the overt realization of the copula in this case is simply motivated by the necessity to make subject AgrS features visible.[8] The finite auxiliary here is then basically of *nominal* (subject agreement) origin.[9] In Arabic, nominal (subject agreement) features are not sufficient to make the verbal auxiliary overt, hence the existence of only verbless present finite sentences (the so-called nominal sentences in the grammatical tradition).[10]

But the overt/covert realization of the nominal copula in French/English (or more generally Romance/Germanic) does not in itself directly explain why the auxiliary is needed to mark finite phi-features of the subject in English or French Perfects or Voices, if we take copular sentences to be structurally distinct from auxiliary ones. I will show that in fact the two instances are not distinct, in spite of auxiliary selection differences.

### 1.3 S/O Agr split and auxiliary selection

Consider again the sentences in (6) and (7). Following early work by Richie Kayne, I assume that the participle there is carrying Part(iciple) (or O(bject)) Agr, and the auxiliary S(ubject) Agr.[11] These two (distinct) Agrs can be made visible when the O raises higher than V, as in (11), where the clitic *les* (for *pommes*), triggers Number and Gender Agr on the participle, and the auxiliary agrees with the high subject:

| | | | | |
|---|---|---|---|---|
| (11) | Je | les | ai | mangées. |
| | I | them-f. | have.1.s. | eaten-f-pl |
| | I have eaten them (f.). | | | |

Note that the S/O Agr distinction or split correlates with another distinction or split in the auxiliary system, namely, the *avoir/have* vs *être/be* split. One feature of this split is that in the *be* configuration, the subject agrees with both the auxiliary and the participle, but in the *have* configuration, it agrees only with the auxiliary, as in the following minimal pair:

| | | | | | |
|---|---|---|---|---|---|
| (12) | Les | tasses | sont | cassées. | |
| | The | cups | are | broken | (-f-pl) |
| (13) | Les | tasses | ont | cassé | (* cassées). |
| | The | cups | have | broken | (*-f-pl). |

In both cases, the cups are understood as "logical" Os, but what matters is that the agreement chain breaks with *avoir*, but not with *être*.[12]

The S/O agreement split (coupled with auxiliary split) is subject to parametric variation. In Ls with a unique copula/auxiliary "be" (the majority, it seems), the agreement chain does not break, even in complex tenses. Consider the following Arabic examples, which illustrate the uniformity of the Agr chain:

| | | | | | |
|---|---|---|---|---|---|
| (14) | kaanat | l-banaat-u | (qad) | katab-na | r-risaalat-a |
| | was-f. | the-girls-nom | (already) | wrote-f.pl. | the-letter-acc |
| | The girls had (already) written the letter. | | | | |
| (15) | kaanat | r-risaalat-u | (qad) | kutiba-t | |
| | was-f. | the-letter-nom | (already) | wrote.pass-f. | |
| | The letter had been (already) written. | | | | |

In this pair, the auxiliary is invariably "be" with the active and passive past perfects alike. At the same time, both the auxiliary and the thematic verb agree with the same structural subject. Fassi Fehri (1993) argues that this configuration obtains through *anaphoric* agreement (and/or anaphoric T for other purposes). At any rate, in both copular and auxiliary sentences, "be" is used, and there is no S/O Agr split. There is only an Agr chain (which includes two AgrSs for the thematic and auxiliary verbs, or for the auxiliary and the participle or adjective). The Agr chain is uniform and controlled by anaphoric subjects.One can then conclude that there is no AgrO in Arabic, and the absence of this split is correlated with the absence of auxiliary split.[13]

If this reasoning is correct, then we have another source for analysis, in addition to the nominality of the copular/auxiliary. The S/O Agr split causes a verbal split (in terms of the auxiliary and the thematic verb), in order to support separately the two Agrs (in fact Agr1 and Agr2). Interestingly enough, the Latin data shows that it is the combination of the two properties (nominal Agr and Agr1/Agr2 complexity) which triggers analysis. Before presenting this case, I will first examine how Temp complexity can yield analytic expressions, through what I call Temp auxiliaries.

## 1.4 Temp auxiliaries

We have seen that the overt realization of the auxiliary in sentences like (8), denoting a state which holds at the present, is not allowed by the grammar. But a version of (8), with the present form of the auxiliary, is possible, provided a different temporal interpretation from the actual present is induced. For example, (16) can express a future state (the equivalent of "The man will be in the house"), or a habitual state (e.g. in the context "The man is usually in the house after 4 o'clock"):[14]

(16) r-rajul-u y-akuun-u fii d-daar-i
the-man-nom is in the-house-gen
(ġadan; ba$^{c}$d-a r-rabi$^{c}$at-i)
(tomorrow; after four)
a. The man will be in the house tomorrow.
b. The man is (usually) in the house after four.

Thus temporal/aspectual specifications like these appear to be marked (compared to the unmarked simple Present), and thus force the auxiliary to become overt, to support these temporal features (cf. Fassi Fehri (1993)).

That specified temporality is behind the use of the verbal auxiliary is further supported by its obligatory occurrence with specified complex tenses, be they perfects or imperfects, active or passive. Thus (17) and (18) illustrate the case of Past Perfect and Future Perfect, respectively:

(17) kaana qad kataba r-risaalat-a lammaa daxal-tu l-qaa$^{c}$at-a
was already wrote the-letter-ac when entered-I the room-acc
He had already written the letter when I entered the room.

(18) sa-yakuunu qad kataba r-risaalat-a ġadan
fut-is already wrote the-letter-ac tomorrow
He will have already written the letter tomorrow.

Note that the occurrence of the modal future prefix *sa-* in (17) is not necessary to force the appearance of the auxiliary; only the future meaning is, as noted earlier.

Analytic complex tense constructions are not limited to Past or future Perfects. They extend to Imperfects. The construction (19) illustrates the case of Past Imperfect, whereas (20) instantiates either a Habitual Imperfect or a Future Imperfect:[15]

(19) kaana y-aktubu r-risaalat-a
was writes the-letter-acc
He was writing the letter.

(20) yakuunu y-aktubu r-rasaaʔil-a fii l-$^{c}$ašiyy-i (ġadan)
is writes the-letters-acc in the-evening (tomorrow)
a. He is usually writing the letters in the evening.
b. He will be writing the letters tomorrow.

These contrasts indicate that (isolating) verbal auxiliaries are clearly playing a role in the realization of complex temporality (provided it is marked). Complex temporality leads then to analysis, and Temp GFs are split on two finite verbs, both of which carrying T marks (i.e. typically prefixing or suffixing Person morphology, for Imperfect and Perfect, respectively), in addition to finite Agr1. Temp auxiliaries are widely spread among Ls, to express T splits (caused by Temp complexity). They are not only found in Germanic/ Romance type Ls, which exhibit the Agr1/Agr2 split, but also in Ls like Arabic (or Semitic), which do not instantiate the latter split. It is worth noting that the T/Agr splits can also be contrasted in the domain of Voice, to account for the different sources of the synthetic/analytic options.[16]

## 1.5 Voice

In this section, I will argue that Voice is also sensitive to the occurrence (or non-occurrence) of Temp (verbal) split or Agr (nominal) split in L. Despite the influential paper of Baker, Johnson and Roberts (1989; BJR), where the Pass morpheme was treated as an argumental (Case absorber) affix generated under I, early parallel work on the matter (see especially Fassi Fehri (1988) for Arabic, Tsimpli (1989) for Modern Greek, and Rivero (1990) for Albanian and Modern Greek) has argued that Pass is a head of aspectual nature, which interacts and clusters with other GFs in the verbo-temporal chain (see also Ouhalla (1991) for motivation of the headedness and verbality of Pass, as well as penetrating criticisms of BJR's analysis). Later work on the matter has acknowledged the existence of Voice as a verbal head projection (see e.g. Kratzer (1996), to be ordered among the components of the aspectuo-temporal-modal chain (see e.g. Cinque (1999)). Very recently, Embick (2000 a and b), who also assumes the aspectual character of the Pass morpheme, resorts to Height of Attachment to derive the various (aspectual) interpretations of the passive participle (eventive, stative, etc.). I will not repeat the arguments for the nature of Voice as temporal, relying on the many references pointing to this direction.

If Voice is verbo-temporal, then it is natural that it selects (or dominates) only VPs (see e.g. Bresnan (1982), Rivero (1990), Ouhalla (1991), Levin and Rapaport (1986)), be modified by adverbs (of Manner; cf. Cinque (1999)), interacts with other aspects (Beedman (1982), Fassi Fehri (1988)), and for our purposes yields synthetics or analysis, depending on other components of the MTAspVoV chain (M for Mood, Vo for Voice). The two triggering factors exploited earlier in active Voice will be equally exploited here, namely Temp and Agr splits.

### 1.5.1 *Arabic and anaphoric Agr*

Consider first Arabic, a simple case. Here Voice in complex tense constructions does not behave differently from that of simple tense cases. For example, in the Passive Past Perfect and the Passive Future Perfect, illustrated in (20) and (21), respectively, Pass occurs on the embedded thematic verb, which is both temporal and finite (compare with (2) above):

(21) kaanat r-risaalat-u (qad) kutiba-t
was-f. the-letter-nom (already) wrote.pass-f.
The letter had been (already) written.

(22) t-akuunu r-risaalat-u (qad) kutiba-t
f.-is the-letter-nom (already) wrote.pass-f.
The letter will be (already) written.

Thus Pass is still synthetically temporal, and the complexity of tenses has no impact on its nature, its position, or the rise of analysis. Note also that the two finite Agrs are identical or anaphoric.

### 1.5.2 *Latin and split Agr*

Classical Latin shares synthetic properties with Arabic in the active verbal system, but differs from it in that analytic auxiliaries arise in the former from complexity created by Passive voice, interacting with tense.

Like Arabic, Latin uses no auxiliary in the Present Perfect. The latter is synthetic, as can be seen in (23):

(23) lauda-vi-t
praise-Perf-he
He has praised.

This synthesis contrasts with Romance analysis in this case (see e.g. French (7) above, or Italian *ha lodato*).

In Latin, the past perfect is also synthetic, although it incorporates (contrary to the Present), a past tense form of the auxiliary *esse:*[17]

(24) lauda-v-er-am
praise-Perf-be-Past.I
I had praised.

Giorgi and Pianesi (1991; = GP), on whose data I am relying, argue that Latin differs from Italian in that it has *verbal* active Perfect (or T2, without Agr2), contrary to the latter in which the Perfect (T2 and Agr2) is *adjectival*. Consequently, T2 in Latin, which does not need a specific Agr2, can be compatible with Agr1. Furthermore, no auxiliary is needed to realize finite features. In the Past Perfect, the auxiliary must appear for temporal reasons, but T2 and T1 are associated with Agr1 (in the absence of Agr2), hence the synthesis.

Latin Perfect Passive, however, behaves differently from the active in that it is periphrastic:

(25) lauda-tus sum
praise-Part.m.s.nom be.Pres.1s
I have been praised.

According to GP (*ibid*), the contrast with the active is in fact expected, given that the needed past or passive participle is adjectival, hence carrying its own Agr (or Agr2). Consequently, the overtness of the auxiliary *esse* becomes necessary, to support (finite) Agr1 and T1.

Our interpretation of the Latin data differs from that of GP's. The fact that the auxiliary used in Latin past perfect is "be", and that it "shares" with the thematic verb the same Agr (or Agr1) is an indication that Latin uses no Agr1/Agr2 split to construct its active Perfect. The auxiliary ("be") can be omitted with the present, but not with complex marked perfects. But although the auxiliary is realized in e.g. Past Perfect, Temp complexity is not sufficient to yield analysis. The emergence of an Agr split (into a non-finite participle Agr2 and a finite Agr1) typically in the Pass Perfect case calls then for an explanation. If the Pass Perfect participle contributes Voice and Aspect, and assuming that Present T is empty (as in verbless or Present Perfect sentences), then the only information which the auxiliary contributes is AgrS (or Agr1), and typically the Pers feature. But the present auxiliary, as we have seen earlier, does not (usually) occur only to support Agr1, as seen with the active Perfect. It must then be that something special is going on with Pass Perfect. Suppose that the latter form is temporally complex, with a sufficient complexity so as to disallow the occurrence of T on it. Then the Pass Perfect complex (with no finite T) will surface with non-finite Agr or Agr2 (i.e. with Nb and Gr features, and without Pers, which is licensed only with finite T). Since a non-copular sentence requires a T/Pers association to be finite, an independent auxiliary is needed in this case. We think then that the analysis is dictated by the degree of (Temp) complexity allowed on the thematic verb, and that the explanation has nothing to do with the participle being verbal or adjectival, as in GP. The Pass participle is normally treated as verbal, typically when it is eventive and (implicitly) agentive. The synthesis/analysis has to do with whether Agr dominates a finite or non-finite Temp, a choice which is available or not, depending on the degree of complexity allowed by L. Thus Latin is forced to use analysis (and Agr split) in the Pass Perfect, but not with Active Perfect, assuming that the former is (temporally) much more complex than the latter. Support for this view comes from the fact that the Pass Imperfect in Latin is also synthetic, presumably because the Imperfect is not marked, and it does not reach the degree of complexity of the Pass Perfect.

Unlike what happens with Latin, Arabic Pass Perfect is synthetic. This difference can be attributed to a difference in the nature of the Pass configuration: the latter is (completely) *temporal* in Arabic (in the sense that it is dominated by other temporal elements, including T), and as such allows more complexity than Latin does. Consequently, Latin resorts to a non-finite Agr (Agr2), to dominate the lower temporal complex allowed, and the remaining Temp GFs, associated with finite Agr1, have to find support on the auxiliary. In contrast, the Arabic Agr being the highest in the tree, after all temporal heads are placed, the Arabic synthetic passive (2) ends up having basically the same form (and internal structure) as the synthetic Latin perfect (23). Note that such an analysis corroborates the view that Voice is part of the Temp chain, and contributes to its complexity. It does not rely on a categorization view of Voice, to derive the differences, a view which is clearly unsupported by the distinct behaviour of Latin Imperfect and Perfect Pass. The following example illustrates the synthetic nature of the Latin Pass Imperfect:[19]

(26) laud-aa-ba-tur
praise-th-past-pass.3.s.
He/she was being praised.

### 1.5.3 *Modern Greek*

Modern Greek (= MG) patterns like Arabic in being synthetic with simple Tense/Voice combinations. Thus the Aorist Definite Past Passive in (27) is synthetic, with an internal *th* affix indicating Pass, and an external one indicating T, Asp, and Agr:[20]

| | | | | | | |
|---|---|---|---|---|---|---|
| (27) | Ta | paidiá | plúthikan | apó | ton | Giánni |
| | the | children | were-washed | by | the | John |

MG differs from Arabic, however, in that its active Present Perfect (in addition to other perfects) is analytic, as in (28):[21]

| | | | |
|---|---|---|---|
| (28) | éxo | yrápsi | |
| | I | have | written |

Likewise, Passive is analytic with the coumpound tenses:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (29) | Ta | paidiá | éXoun | plúthei | apó | ton | Giánni |
| | the | children | have | been-washed | by | the | John |

As proposed by Rivero (*ibid*), *th* in (29) occurs on the main V with the same role, and the auxiliary *éXoun* fails to reflect Voice. Assuming a hierarchical structure T>Asp>Vo>V, she argues that Voice is always an affix head immediately above VP. In simple tenses, Pass attaches to V, and the complex to Asp, then to T. But in perfect tenses, Asp is the stem *eX*- "have", rather than an affix. The aspectual Aux raises (separately) to T (and Agr), and the V to Pass (only). However, Rivero does not provide an explanation for the split.

In my view, these facts can be interpreted as follows. Pass is (totally) verbo-temporal in simple tenses because no complexity arises there, in terms of T or Agr, combined with Voice. The existence of an analytic Present Perfect is a diagnostic indicating that MG exhibits the Agr split (and/or complexity) even with actives. As for complex tenses, MG is not different from Arabic in being analytic. Furthermore, the Pass affix in MG behaves like the Arabic one in that it does not play any role in analysis, unlike what happens e.g. in Latin.

### 1.5.4 *Albanian*

As described by Rivero (*ibid*), Albanian represents an original case of variation in Pass morpho-syntax, and typically how it interacts with structural complexity. The Pass morpheme appears to be *mobile*, and it occurs as an affix, a clitic, or part of Aux, depending on contexts. The first two cases are synthetic, in that Pass is part of the thematic V compound, while the last case is analytic, appearing on the auxiliary.

First, consider the simple past passive (or non-active; Nact) in (30):

(30) La-he-sh-a
wash-Nact-Past-1s.
I was washed.

The internal structure of this construction is (essentially) similar to that of MG (27), except that no Aspect projection can be justified in this case. Pass is adjacent to (or dominates) V, and is dominated by T. This is an instance of affixal Pass.

Second, when internal Mood such as the Optative (for wishes) is used, it occupies the position of Voice, and Pass is displaced to a clitic position, as in (31):

(31) U la-fsh-a
Nact wash-Opt.Pres-1s.
May I be washed.

Third, Perfect tenses indicate Voice in the auxiliary, regardless of Mood. Thus the auxiliary is *kam* "have" in the active Future Perfect (32), and *jam* "have been" in the passive (33):

(32) Do të kam läre
Fut have.Pres.1s. washed
I will have washed (something).

(33) Do të jam läre
Fut have+been.Pres.1s. washed
I will have been washed.

In order to account for the *mobility* of Pass, Rivero proposes that the Albanian thematic verb be restricted to carrying three affixes only. In the affix case, the three are: Vo, T, and Agr, in this order, simple Albanian verbs normally lacking aspectual distinctions (except the Aorist Definite Past, which behaves like Mood, and leads to clitic Pass; see her example (13)). In the clitic case, she assumes that Mood, Aspect, and Voice are generated under the same node dominating VP. When Mood or Aspect are realized, Voice is not part of the three affixes, and it is raised to a clitic position. As for the fusion with Aux, Rivero explains that that operation makes it adjacent to V, the essential requirement.[22]

It is clear then that the interpretation of the facts about synthesis and analysis in Albanian can be understood first in terms of complexity. They are essentially like those of MG: synthesis with simple tenses, and analysis with compound Perfect tenses. The existence of Present Perfect indicates the existence of an Agr split. This split is reinforced by temporal complexity. The mobility of Pass, typically when it fuses with Aux, if correct, opens up an original option for the expression of Voice, built also on complexity. But Voice is not behind analysis. Yet the fusion of Voice with Aspect and T provides further support for the hypothesis that Voice belongs to the aspectuo-temporal chain.

### 1.5.5 *Moroccan Arabic*

I will use Moroccan Arabic (= MA) as a prototypical example of spoken Arabic dialects to analyze Pass voice and M-T-Asp interactions (given that differences may be taken as of the microvariation kind). The MA Temp chain, despite appearences, is very similar to that in Standard Arabic (=SA) described above, except that MA, like other Arabic dialects, has lost its apophonic marking of Pass, due to an extensive loss of stem vowels. Consequently, the morphemes [*n*-] or [*t*-] (or their combination [*nt*-, *tn*-], depending on regional or individual microvariation), which were used to express a kind of detransitivization with the meaning of inchoative, reflexive, or reciprocal, but distinct from passivization (expressed via apophony), were extended in MA (and other dialects) to express also Pass (or a medio-passive). So the difference in morphological expression between the inchoative (34) and the passive (35) in SA turns out to be expressed by the same [*t*-] prefix in MA (36):

(34) ta-kassar-a l-kaʔs-u
inch-broke.intens-3.s.m. the-glass-nom
The glass broke into pieces (came to break).

(35) kussir-a l-kaʔs-u
broke.intens.Pass-3.s.m. the-glass-nom
The glass was broken into pieces (by someone).

(36) t-hərrəs l-kaas
Pass-broke.3.s.m. the-glass-nom
a. The glass was broken.
b. The glass came to break.

One essential difference between the two members of the SA pair is that no implicit agent is inferred in (34), but there must be one in (35). But in (36), the difference is neutralized, and in fact, the (b) interpretation appears to be the most natural one, unless an extra PP-like by-phrase is added. Also, the morphology does not lead to a different conjugation, and appears to behave like the active SA morphology. But before going further with the non-active verb, let us see how the inflectional morphology is built.

MA does not differ essentially from SA in this respect. Consider the following two pairs of verbal constructions, the first one being active indicative, and the second one being non-active indicative:

(37) ktəb-t l-ebra
wrote-I the letter
a. I wrote the letter.
b. I have written the letter.

(38) ta-y-kətb-uu l-ebra
indic-3-write-pl.m. the letter
a. They are writing the letter.
b. They write the letter.

As in SA, the past in (37) is marked by a (continous) suffix indicating Pers, Nb, and Gr ([- *t*] there). The form expresses also the Present Perfect, without auxiliary use, while Past and Future Perfects resort to auxiliaries. In the present (38), tense is expressed by a discontinuous affix, with a prefixing Pers, and a suffixing Nb and Gr, as in SA. The form is also ambiguous between a progressive reading and a habitual reading, among other potential meanings. The first prefix [*t*-] there expresses Indicative (or Realis) Mood, as opposed to Subjunctive or Jussive (or more generally Irrealis), marked by its absence. The [*t*-] position can also be occupied by the prefix [*ġa*-] (a short form of *ġaadi*), which I take to be a modal for future. This is a further similarity with SA, which puts modal affixes like the future [*sa*-] in this position. Note that [*t*-] cannot be analyzed as progressive (or durative) Asp marker, since it can appear with stative verbs:[23]

(39) ta-n-ᶜref
indic-I-know
I know.

(40) ta-y-kuun hna
indic-3-is here
He is usually here.

Clearly, neither (39) nor (40) have progressive reading. In (41), without [ta-], the interpretation is hypothetical (Irrealis):

(41) y-kuun hna
3-is here
He might be here.

The [*t*-] prefixed form is then expressing a marked Realis Mood, conditioning the actual or habitual Present, whereas the bare prefixed form expresses a non-real Present.[24]

Consider now the passive case:

(42) t-əḥrəq l-xubz
Pass-burned-3f.s. the-bread
a. The bread was burned.
b. The bread has been burned.
c. The bread has burned.

(43) ta-y-t-əḥrəq l-xubz
indic-3-Pass-burn-3.s.m. the bread
a. The bread is burned.
b. The bread burns.
c. The bread is burning.

Leaving the details of ambiguities aside, the order of prefixing morphemes in (43) is the following: M-T-Vo-V, with a Pass prefix being the closest to V, and lower than T/Asp, an ordering also found in SA. I assume tentatively that in the reflexive/reciprocal/inchoative interpretations, [t-] is prefixed to the root (before category labelling), as in SA, and in the medio-passive case prefixed to the stem (after category labelling). This difference in attachment height mirrors the clearly overt morphological distinction found in SA, where the first class of affixes (which are consonantal) attach to the root (to form a complex root), while the vocalic passive infix contributes to category stem formation. At any rate, the passive affix is found over V, and below T and M, in the M-T-(Asp)-Vo-V chain. It is therefore part of a non-split temporal chain, and hence synthetic.

## 2. Formal complexity and categorization

In the previous section, I have argued that analysis results from two types of splits:

a. *Agr (nominal) split,* found typically in Ls that have either (a) an analytic present Perfect (e.g. Romance and Germanic), or (b) an analytic Pass Perfect (e.g. Classical Latin) or (c) both (e.g. Germanic or Romance).
b. *Temporal (verbal) split,* found in Ls like SA (and Arabic dialects) which (a) do not have an analytic active present Perfect nor (b) an analityc Pass Perfect. They do exhibit analysis, however, in complex tenses (in e.g. Past or Future Perfect tenses).
c. Voice (Pass) has been argued to be part of the Temp chain, and it triggers analysis by interacting with Temp complexity (e.g. in Latin).

The source or cause of analysis assumed here can be called *Formal Complexity*, that is complexity in morphosyntactic expressions of GFs (= FC), and typically formal Temp complexity, conjoined with properties of (formal) Agr. FC makes predictions about the kind of analytic types to be found in Ls, to the exclusion of others, and the nature of limitations to be expected on choices, such as e.g. the Formal Disjunctivity stated in (1) above. Other approaches provided in the literature (including those based on Cognitive complexity or pattern change) do not make the right predictions.[25] Before returning to this issue in Subsection 2.3, I first address further relevant questions that remain to be answered.

## 2.1 Further analytic and synthetic questions

### 2.1.1 *Pass and additional complexity*

Why there seem to be no Ls that exhibit only the Aa property, without exhibiting the Ab property (i.e. Ls which would be the reverse of Classical Latin)? The premium answer seems to be that Pass adds an extra level of complexity to the active Present Perfect. So if the active is already analytic, its passive counterpart should also be so (but the reverse is not true). The added complexity of Pass appears to be natural if Pass is marked, compared to the active, and it adds extra temporal specification (two-degree specification here).

But if Pass is temporal, as I have argued, why should it trigger analysis in Classical Latin Pass Present Perfect (26), but not in active Past Perfect (25), the two cases appearing to exhibit the same degree of complexity? Various answers are possible here, all of which would lead to add an extra factor that would either make (26) much more complex than (25), i.e. 3-level complex, or make the extra factor contributing to increase general complexity. I will adopt tentatively the former solution.[26] Then the T-chain would be associated with one (finite) Agr in (25), but two Agrs in (26). In Ls that exhibit the *have/be* split in auxiliaries, the Pass non-Pass distinction is usually correlated with auxiliary selection. In actives, the auxiliary (*have*) carries the Agr of the external subject (or Agr1), whereas the perfect participle is carrying an (abstract) Agr2, which is disjunct from Agr1 (and does not agree with it). With the Pass (Perfect), however, Agr2 must be anaphoric with Agr1 (and the auxiliary used is *be*). But given the relevant complexity in Classical Latin, and also the fact that there is only a single auxiliary (*be*), the only way to deal with Pass/Temp complexity is to introduce a distinct Agr (or Agr2) for the passive participle subject (the underlying object of the thematic verb), which cannot be directly associated with Agr1 (normally linked to the finite active subject). Moreover, anaphoricity between Agr1 and Agr2 is established, given that the auxiliary is *be*, and the subject of the participle is also the subject of the finite (auxiliary) verb.

### 2.1.2 *Two finite Agrs*

If Arabic Present Perfect is synthetic, why can't its Past Perfect be, as in Latin? I would claim that Arabic does not behave like Latin in this case presumably because of the conjunction of two factors. First, Arabic has no distinction (in its inflectional inventory) between finite and non-finite verbal morphology (such as participal or infinitive). All temporal verbs are finite (as exemplified above), and deverbal adjectives have a different morphology, as we will see below. In order to form a complex tense, Arabic resorts only to the tensed auxiliary "be" (*kaana*), and to a finite verb which carries the same Agr form, the latter being anaphoric to the upper Agr. In other words, Agr1 and Agr2, which enter in the formation of complex tenses in this language are identical. Second, Agr morphology is hardly dissociable from T because it is essentially the positioning of the Agr morpheme on the verb (as prefix or suffix) which indicates past/non-past (T1) or perfect/imperfect (T2).

### 2.1.3 *Ancient Greek as fully synthetic*

Are there Ls which are completely synthetic, in the sense that they exhibit no Agr split, and no Temporal split? Ancient Greek (= AG), as described by Ackema (1999), appears to provide an instance of such a language.

Consider Present active forms, where the Perfect inflection and the T/Agr inflection are all affixed to the thematic verb *lùo* "loosen":

(44)

| | Sg | Pl |
|---|---|---|
| | 1. léluka | lélùkamen |
| | 2. lélukas | lélùkate |
| | 3. léluke | lélùkaasi |

In the conjugation of e.g. *lélùkamen*, Perfect morphology consists of reduplicating the stem (of the non-passive voice), joined by a *k*- suffix; T and Agr then follow. The active Perfect behaves essentially like simple active tenses, except for the (internal) introduction of Perfect morphology in adjacency with the verbal stem.

The paradigm for the Perfect passive remains synthetic, although three (instead of two positions) are filled, with Pass (or the medio-passive) generated in a position higher than Perfect, and adjacent to I (or T/Agr):[27]

(45)

| | Sg | Pl |
|---|---|---|
| 1. | lélumai | lélùmetha |
| 2. | lélusai | lélusthe |
| 3. | lélutai | léluntai |

If this is correct, then the double complexity induced both by Pass and Perfect in AG does not lead to a Split (of Agr or Temp).

### **2.2** Reanalysis as the source of analytic Pass or Perfect

A number of studies agree that a reanalysis took place in Germanic and Romance to yield periphrastic Perfect passives, and then generalized to Perfect actives (and non-perfects as well).[28] Ackema (among others) argues that deflexion (or the loss of inflectional endings), as well as the number of mismatches and ambiguous mappings between the number of affix positions in morphosyntax and the number of morphonological affixes showing up on the verb make parsing opaque, and hence yields to analysis (through reanalysis). Taking synthesis to be a parameter in terms of the number of positions that a verb takes (with a maximum of three, Spec, Comp, and Adjunct), he then postulates a possible decrease in synthesis, with the Spec being the last position to undergo analysis. Deflexion is then taken to be the cause of a reanalysis of an already existing construction in the language. The target in the periphrastic passive case is a copular structure, consisting of the copula *be* and the deverbal adjectival passive. The (reanalyzed) adjective as participle is a true adjective by standard tests. Thus in *victus est*, the original meaning is that the subject is in a conquered state. When the construction was reanalyzed as Perfect passive, the new meaning became "it has been conquered by someone". The new passive does not have a distinct perfect inflection (like synthetic perfects have), the new auxiliary *be* being construed as inherently perfect, with the completion of the action (expressed by the adjective's verbal base) implied. The original structure is not lost, and a structural ambiguity is introduced in the syntax. If deflexion persists, then only one position is left on the verb. Consequently, after the passive becomes analytic, synthetic forms of the non-perfect passive and active perfects start disappearing, and the same deverbal adjective construction is used, with a usual combination with *have* for the Present Perfect. Moreover, the reanalysis introducing the periphrastic verbal structures also introduces the present day ambiguity between the latter and their adjective ancestors.

One claimed advantage of the analysis/reanalysis proposed is that in Germanic and Romance, Pass and Perfect participles are predicted to be identical in form (as well as their deverbal adjective counterparts), whereas synthetic Pass and Perfect have different forms. The author rightly criticizes alternative approaches like that of Ouhalla (*ibid*) or GP (*ibid*), based on categorial parametrization of Perfect or Pass. For example, GP analyze T2 as *verbal* in the active, and *adjectiva*l in the passive, contrary to the evidence that Pass can be verbal as well. The same objection applies to Ouhalla's *nominal* feature of Pass in analytic cases. Furthermore, these analyses (which Ackema terms "head movement analyses"), do not predict any relation between richnesss of inflection and the synthetic/periphrastic distinction. In sum, the analysis proposed establishes a direct link between deflexion and replacement of synthesis by periphrasis. In the criticized structural analyses, the distinction is characterized via the category of GFs, but it is unclear why there should be a relation between deflexion and the categorial feature change in Pass, Asp, or T.

Embick (2000a) also adopts a head movement approach to the synthetic/analytic distinction in Latin Perfect. Given a unified structure for both cases and a list of GFs, including v, Asp (under which he places Perf or Imperf), Pass (which has no determinate position), and T, he then simply stipulates the condition (46) to prevent synthetic Pass Perf (= his (29)):

(46) [Perf] Asp does not move to T when [pass] is present.

But this stipulated statement can hardly be used to predict the analytic types available across Ls, or account for how the variation arises. For example, (46) does not apply to Classical Greek, and we do not know why. On the other hand, one could change Perf with Imperf and obtain an analytic type which does not exist, or change passive with non-passive, to make an equally wrong prediction.

## 2.3 A Splitting Analysis (of Temp and Agr categories)

Although valuable on certain points and criticisms, Ackema's analysis suffers from a number of drawbacks. Assuming that the analysis/reanalysis proposed accounts correctly for the patterns it deals with in essentially Romance, Germanic, Latin, and Greek, (a) it is also based on a *category change* explanation, (b) it is limited to a *single type of Ls* where the same path is used to move from synthesis to analysis, without covering a wide range of distinct crosslinguistic paths, (c) it does not take care of *peculiarities* in Ls that make a simple formulation of the Synthesis Parameter in terms of the hierarchy and number of the (argument or adjunct) affixes visible on the verb workable, and (d) it does not address the question of how predicate affixes, inchoatives, causatives, reflexives, or modal affixes (to cite few cases) fit into the picture. For example, deflexion cannot explain why English developed an analytic Present imperfective, whereas French does not, hence the contrast in (47):

(47) a. Jean mange.
b. John is eating.

The same is true of English Past progressive, as opposed to the synthetic French "imparfait". The lack of analytic imperfective in French is also found in the passive patterns, as we will see. Likewise, deflexion can hardly be behind the replacement of punctual simple Past in French "passé simple" with an analytic ambiguous "passé composé", serving also as general past. The reanalysis simply stipulates a category change, independently of the other morpho-syntactic properties of Ls. If deflexion were sufficient for reanalysis, then we expect MA (in which apophonic Pass morphology was lost, due to a general loss of internal stem vowels) to resort to a reanalysis of the adjectival passive construction along similar lines. The reanalysis has simply extended the use of the [*t*-] inchoative affix to Pass, keeping verbal Pass and adjectival Pass separate. Finally, it is not clear why Ls like Classical Latin do not undergo a reanalysis like Italian in the active Perfect.

Keeping these objections in mind, we can now go back to the link we have established between morphosyntactic complexity and its morphonological realization. Limitations on the mappings are language specific, just as the resort to analysis is. We do not pretend that FC leads automatically to analysis, but we assume that peculiarities of the language aside, it does. For example, Voice occurs as verbal or adjectival in a number of Ls, but the variation is whether it occurs in finite or non-finite (participial) verbs. The option is then whether Voice is dominated by a T/Asp head associated with a finite Agr (or Agr1), or whether it is dominated by a non-finite T2/Agr2. The former option is found in synthetic passives, and the latter in analytic ones. Assuming then the role of complexity in leading to splitting (or fissioning) in syntax, let us designate the analysis by what we take to be its key feature, and call it *Splitting Analy-*

*sis*, and let us restate its parameters (with respect to Agr-T-Vo complex chains) as follows:

(48) *Splitting Analysis*

a. Split Agr into finite Agr1 and non-finite Agr2 if Aux is of nominal nature.
b. Split the T chain into T1 and T2 if Aux is verbal.
c. Split a finite Agr/T chain into a finite Agr/T1 and non-finite Agr/T2 (and attach Pass to non-finite Agr/T2).

These split operations take into account the peculiarities of Ls, as we have seen, with the absence of the Present copula in Arabic and Latin (the negative instance of (48a), the non-separability of T and Agr in SA (a case of (48b)), as well as the high level of analysis in Germanic and Romance (a case of (48c)). In the next section, I address some other interactions betwen L peculiarities and these splits, as well as the role played by Height of Attachment of a Temp or Agr affix (instead of categorization) in deriving differences inside a single L, or between Ls.

## 3. Peculiarities and structural heights

### 3.1 Imperfective Passives

Arabic imperfective passives exhibit significant properties, and like their active counterparts, undergo no limitations on their formation or meaning. This is not the case for example in French, in which there is no such passives. On the other hand, English does have such imperfectives, although not with exactly the same distribution of meanings.

Thus Arabic forms both Past and Present passive imperfectives, as exemplified by (49) and (50), respectively:

(49) kaana l-xitaab-u y-uktabu
was the-speech-nom 3-writes-pass
The speech was being written.

(50) yakuunu l-xitaab-u y-uktabu
is the-speech-nom 3-writes-pass
a. The speech will be being written.
b. The speech is usually written.

Consider now how the Past or Future progressive passives are expressed in French. French has no grammatical way to express these meanings, and must resort to lexical periphrasis, as illustrated by (51) and (52), respectively:

(51) Le discours était en train d'être écrit.

(52) Le discours sera en train d'être écrit.

This lack of grammatical parallelism between French and SA can be traced back to the fact that French (a) lacks a progressive participle (which can enter a compound Temp chain, as in English), and (b) lacks the possibility of combining two finite verbs, the lower of which can be interpreted as imperfective/progressive (a possibility available in SA).

Let us turn then to English. The English translation of SA (49) is acceptable. Then there is no significant difference with SA, except that the embedded passive is a synthetic verb in SA, but an analytic participle in English. With regard to (50), the use of the expression "in the process of" appears to make the acceptability easier, but the sentence is nonetheless acceptable. Now, acceptability apart, the two sentences do not share the same meanings.[29] SA (50) can only mean a future passive progressive. But a real present imperfective passive interpretation is available in English:

(53) The speech is being written (now).

Likewise, SA (50) is ambiguously habitual, but its English present counterpart is not. These peculiarities recall the temporal conditions under which the present of the auxiliary and/or copula may be overtly realized in SA, English, or French. Moreover, non-available interpretations are attributable to the limitations arising from the absence of the imperfective passive (in the French morphological inventory), or the limitations on non-finite (past participle) Passive interpretations (in English and French). The availability of finite verbal Voice and finite imperfective in SA appear to allow more combinations.

### 3.2 Verbal and adjectival voices

Another striking difference between Arabic, English, and French is how Voice combines with verbs or adjectives to express e.g. Passive, and how the eventive/stative reading split in passives is captured in their various grammars. Recall that in the early literature on English passives (and Ls of similar type like Romance), the debate was first centered around the *category* of the passive participle. Thus the eventive agentive interpretation of (54), for example, given in (54a), was attributed to the *verbality* of the participle (the verb undergoing no category change there), while its stative non-agentive interpretation, given in (54b), was supposed to follow from its *adjectival* nature, derived through a category-changing operation (cf. in particular Wasow (1977) and Bresnan (1982); cf. also Levin and Rapaport (1986)):[30]

(54) The glass was broken.
  a. Someone broke the glass.
  b. The glass was in a state of having been broken.

But as argued e.g. by Embick (2000b), relying on a comparison of English, Chichewa and Creek passive expression, category is not the deciding factor. For example, there is no direct correlation between the syntax/semantics of stativization and the surfacing category.[31] It must then be that it is the distinct syntactico-semantic properties of the aspectual category heading the participle which make the difference. Thus Embick resorts to Height of Attachment of the aspectual head to derive the difference (in lines with Kratzer (1996) and Marantz (2000), among others). With statives, Asp is directly attached to the Root, and with eventives, it dominates small *v*. If there is no *v*, then no event interpretation is available. In sum, Embick appeals to structure heights to derive the ambiguous behaviour of an otherwise identical participial passive affix.[32]

But note that Arabic does not exhibit ambiguities in Voice expression, similar to those found in English or Romance. This is so because Arabic resorts to distinct divisions in Voice expressions. The essential dichotomy used here can be identified as *finite* Voice, found with temporal verbs, and *non-finite* Voice, found with atemporal adjectives. To get directly to the distinguished flavour of the picture, observe that the two passive interpretations in (54) are expressed by two distinct constructions, (55) and (56), respectively:

(55) kusir-a l-kaʔs-u
  broke.pass.-past.3 the-glass-nom
  The glass was broken.

(56) l-kaʔs-u ma-ksuur-un
  the-glass-nom broken-nom
  The glass is broken.

Here, there is no room for ambiguity. In (55), the passive vocalic morphology is internal to the stem of the finite verb, which is marked for T and Agr, and which is both agentive and eventive (with the implicit agent, assessed by standard tests, like agent oriented adverbs and control structures). But in (56), the passive has adjectival pattern morphology, it receives Case, and whatever Agr adjectives receive. Moreover, it is stative like any other adjectives. Clearly then, the passive verb in (55) is dominated by v and T (among other functional projections), but the one in (56) is dominated by A (or a small a), and whatever stative aspectual or functional heads are found in the structure of adjectival phrases (cf. Fassi Fehri (1988, 1993) for detailed proposals). Then if Pass (or Voice) is a functional head (as proposed e.g. in Fassi Fehri (1988) and Kratzer (1996)), then Pass dominates v, V, or R(oot) if adjective is internally stative, which are in turn dominated by Asp and T.

A close examination of deverbal SA passive and active adjectives suggests further refinements (cf. Fassi Fehri (1993) for detail). Consider the active and passive adjectives exemplified in (57) and (58), respectively:

| | | | | |
|---|---|---|---|---|
| (57) | r-rajul-u | maaniḥ-un | l-hadiyyat-a | (ġad-an, *ʔamsi) |
| | the-man-nom | giving-nom | the-gift-acc | (tomorrow, *yesterday) |
| | The man is giving the gift (tomorrow, *yesterday). | | | |
| (58) | r-rajul-u | mamnuuḥ-un | l-hadiyyat-a | (ġad-an, *ʔamsi) |
| | the-man-nom | given-nom | the-gift-acc | (tomorrow, *yesterday) |
| | The man is given the gift. | | | |

In both sentences, the adjective is (internally) eventive, assigns accusative Case to its object, and the event is temporally interpreted as taking place either in the present, or the future, as shown by the adverbial tests. But T is not overtly realized. These properties appear to mirror those of progressive and passive participles in English, auxiliary realization aside:

(59) The man is giving the gift.

(60) The man is given the gift.

The question that arises then is the following: what is the difference between the translation of (59) or (60) by active or passive (internally eventive) adjectives such as those in (57) or (58), and its translation by finite verbs (in the present form), as illustrated by the following pair:

| | | | | |
|---|---|---|---|---|
| (61) | r-rajul-u | y-amnaḥ-u | l-hadiyyat-a | (ġad-an, *ʔamsi) |
| | the-man-nom | 3-gives | the-gift-acc | (tomorrow, *yesterday) |
| | The man is giving the gift (tomorrow, *yesterday). | | | |
| (62) | r-rajul-u | y-u-mnaḥ-u | l-hadiyyat-a | (ġad-an, *ʔamsi) |
| | the-man-nom | 3-pass-gives | the-gift-acc | (tomorrow, *yesterday) |
| | The man is given the gift (tomorrow, *yesterday). | | | |

It appears at first glance that the two constructions are equivalent, but upon closer examination, they are not. One striking difference between the two cases is that only verbs, but not adjectives, take part in complex tense formation, as shown by the following contrasting pair brought up in Fassi Fehri (*ibid*):

| | | | | | | |
|---|---|---|---|---|---|---|
| (63) | kaana | r-rajul-u | y-ajlis-u | $^{c}$indamaa | daxal-tu | l-qaa$^{c}$at-a |
| | was | the-man-nom | 3-sits | when | entered-I | the-room-acc |

The man was (in the process of) sitting when I entered the room.

| | | | | | | |
|---|---|---|---|---|---|---|
| (64) | kaana | r-rajul-u | jaalis-an | $^{c}$indamaa | daxal-tu | l-qaa$^{c}$at-a |
| | was | the-man-nom | sitting-acc | when | entered-I | the-room-acc |

The man was sitting when I entered the room.

As observed there, the sentence (63) is ambiguous in a way that (64) is not. In the latter, the adjective is stative, and when I entered, the man was already (in the state of being) sitting. But (63) can have a reading in which he has not reached that state yet, i.e. a progressive reading. Clearly then the adjectives in embedded contexts are necessarily stative, whereas verbs are not. Likewise, the verbal passive (62) is progressive, whereas the adjectival (58) is not. These contrasts indicate that there is a clear separation in Arabic between finite verbs and (deverbal) active or passive adjectives. The former are temporal, whereas the latter are not. Furthermore, the fact that those adjectives do not take part in complex tense combinations indicates clearly that they cannot count as verbal participles, i.e. they possess no "mixed" or variable behaviour category (paralleling that of English or Romance), which would enable them to alternate between a category which does take part in the temporal system (i.e. participles dominated by Perfect or T2) and a category which does not take part in it (i.e. stative adjectives).[33]

## 3.3 Multiple functions across heights

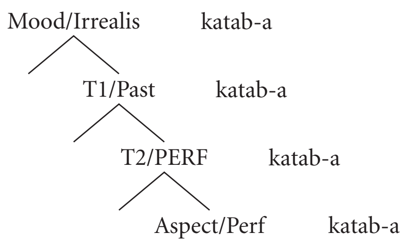
It is widely spread property of Ls that a morphological form is often ambiguously used to express different GFs, or the same GF at different levels of structure. As I have amply argued elsewhere (see Fassi Fehri (1993), and 1996b/2000, in particular), verbal morphology encodes not only Tense, Mood, and Aspect (= MTA there), but also Voice and Agr. These functional traits (or categories) are organized hierarchically, according to (65), leaving Agr aside:

(65) Mood > T1 > T2 > Aspect > Voice

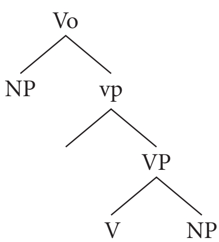
T1 stands for the Past/non-Past distinction, T2 for the Perfect/Imperfect, Aspect for Perfective/Imperfective. Unlike the Present/Imperfect prefixed form, which marks Mood, the Past/Perfect suffixed form does not. However, this form does have modal uses, in wishes, or hypothetical situations, as illustrated by the following sentence:

(66) raḥima-ka l-laah-u
blessed-you Allah
May Allah bless you!

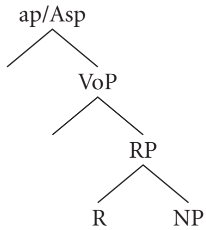
These ambiguous uses of suffixed temporal morphology can be accounted for via either movement or appropriate placement of the same synthetic word under the relevant category. The following gives a rough sketch of how multifunctionality can be implemented (I use the word *katab-a* "wrote" as an instance for repesenting this ambiguity):

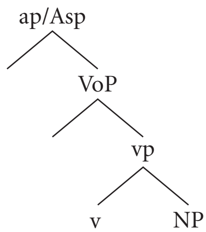
(67)


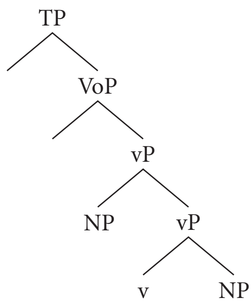

The lower part of the tree, which is irrelevant for the point, is the following:

(68)



As for Voice contrasts, they can be represented in the same way. Three cases are relevant here: (a) the completely stative adjectival passive, which can be instantiated by (56), the (internally) eventive adjectival passive in (58), although with no implicit agent, and the eventive agentive verbal passive in (55). Following the lines of Marantz (2000) and Embick (2000b) in particular, these Voice contrasts can be represented roughly as follows (R for Root):

(69)



(70)



(71)



These structures are given for the sake of illustration. In the totally stative (69), Vo is heading RootP (before category formation), and the category (with its aspectual properties) dominates Vo. In the partly eventive (70), Vo dominates an unaccusative vP, with an internal argument only. The small v makes it eventive, but no Agent can be implied. The Asp/ap converts it to an adjective externally. In the verbal eventive and agentive (71), the vP is transitive, dominated by Vo, but the latter is dominated by TP, which makes the verb finite.

## 4. Summary and conclusion

In this article, I have shown how Temporal and Voice expressions in a number of Ls distribute with respect to the synthetic/analytic continuum, and what kinds of limitations and constraints operate across Ls, and inside a particular L, taking into account its peculiarities. *Formal Complexity* (FC) has been argued to be behind changes from synthetic to analytic expression, once a degree level of complexity is reached in the Temp chain. Appropriate predictions are then made about potential analytic types to be found. *Splitting Analysis* operates on Temp chains, as well as on Agrs, to create two separate, though anaphoric chains. Variation splits are argued to depend on the nominal or temporal nature of the auxiliary features in L, the existence (or non-existence) of non-finite (temporal) Agr in L, and the ability of Vo or Pass to be dominated (or not) by a Temp or an Agr head (see the parameters in (48)). Finally, *Formal Disjunctivity* has been proposed to limit the choices in Ls to either synthetic or analytic expression of GF in L, but not both, hence making the choices typically asymmetrical.

## Notes

* This article is dedicated to the memory of Ken Hale, a great linguist and a great friend. For useful discussions, I would like to thank Alec Marantz, Morris Halle, Jim Higginbotham, Anna Maria Di Sciullo, Sabine Iatridou, Sylvain Bromberger, Abderrazak Tourabi, and for help with the Ancient Greek data Donca Steriade. I also thank the audiences at the Semitic Conference at USC (May 2001) and the Asymmetry Conference at UQAM (May 2001), as well as an anonymous reviewer of the Proceedings for helpful remarks. The usual disclaimers apply.

**1.** LBM theory is claimed to have *arisen* in order to solve the problems of a number of asymmetric mappings of sound and meaning (at the morphological level), which cannot be solved in morpheme-based grammars (= MBG). The latter take morphemes to be signs. Among those problems are the following:

i. Mattews (1972)'s "cumulative exponence", i.e. the fact that a single form expresses multiple GFs simultaneously (e.g. Latin verbal *-o* expressing: 1.s.pres.indic.);

ii. M's "extended exponence", with multiple forms expressing a single GF (English "*be*" + *–ing* making the progressive);

iii. Multiple functionality (English *-ing* with v, n, a GFs)

iv. Multiple expression of a single GF (English *-ity, -ness, -ce* marking deadjectival transposition).

To these asymmetries, null and empty morphologies should be added. The Separation Hypothesis amounts to separate the terminal elements in syntax from their phonological realization. For detail, see Beard (2001).

**2.** DM combines insights of both LBM and MBG. In DM, (i) syntactic features and phonological features form separate sets, (ii) affixes are lexical items , which combine with lexemes or other affixes to form complex words, (iii) vocabulary items compete for insertion into syntactic nodes, and (iv) morphological structure is a separate component of the grammar.

Later development of DM has modified some of these assumptions, namely (iv). For details, see in particular recent work by Marantz (2000), among others.

**3.** Comrie (1985), for example, is rather pessimistic about the utility of these typologies, typically for the purpose of correlating with other non-morphological properties of L, when he writes:"Although morphological typology does serve the useful purpose of presenting an overview of the morphological structure type of a language, it remains unclear whether it can be considered a significant typological parameter [...] in the sense of *correlating with other parameters outside morphology* [our underlining; FF]" (p. 52).

**4.** These morphological types has been proposed originally by Friedrich von Shlegel (1808), August Wilhem von Shlegel (1818), as well as Wilhem von Humboldt. See Bussman (1996) for definitions as well as references.

**5.** I use "formal" as a cover term for both morphological and syntactic relations, which deal with form, as opposed to meaning or semantics. FD parallels similar asymmetric constraints proposed namely by Di Sciullo (for recent instances, see Di Sciullo (1999)). Apparent counterexamples to (1) appear to be numerous at first sight, but at a closer examination, the disjunctivity will hopefully be confirmed. For example, during discussion at UQAM, Jean-Yves Pollock has raised the case of the French future pair *Je vais dormir* and *Je dormirai* as equivalent, but Denis Bouchard has objected to the judgement. Edwin Williams has put forth the case of periphrastic and synthetic comparative adjectives like *much more clever* and *cleverer*, which apppear to be problematic at first sight. Presumably, there are subtle differences between these pairs that I will not go into here. Another very discussed case in the literature is the difference between "lexical", "morphological", and "periphrastic" causative doublets or triplets. My guess is that the problem of these apparent counterexamples can be solved in a major part through height of attachment of the GF in the structure, from which differences in meanings result. Furthermore, L may be at some intermediate stage of change where doublets appear to be equivalent, before the change is stabilized.

**6.** On the role of nominality in forming complex tenses, see Fassi Fehri (1996a).

**7.** I leave for the moment the contrasts in the past tenses, which are more complex, and do not distribute in the same way. Thus French contrasts with both Arabic and English in requiring an auxiliary for the expression of its simple Past:

i. Jean a mangé la pomme hier.

ii. kataba r-risaalat-a ʔamsi
wrote the-letter-acc yesterday
He wrote the letter yesterday.

iii. John wrote the letter yesterday.

French has also a synthetic imperfect Past (the "imparfait"), which is comparable to the (analytic) past progressive in English:

iv. Jean mangeait la pomme hier.
John was eating an apple yesterday.

I return to some of these contrasts in Section 2.

**8.** On the absence of Present tense morphology (in English), see in particular Kayne (2000, Ch. 10, originally published in 1989), and Enç (1991). For a different view, see Fassi Fehri (1993) and Guéron (1995), among others.

**9.** I assume that copula and auxiliary "be" do not differ significantly as regards the conditions on their "visibility". The copular sentence presumably differs from the non-copular one in that the latter contains "be" plus a temporal thematic participle. The auxiliary qualifies as nominal because it carries only features of the (finite) Subject Agr (= Agr1), T1 being null. I assume that participles are dominated by a nominal and/or adjectival (non-finite) Agr (= Agr2), although they may be also headed by a null abstract Aspect or T head (= T2, needed for appropriate interpretation). Participles may then be taken as "weakly" temporal, because they lack overt temporal marks in all contexts, apart from Voice. See Section 3 below.

**10.** For the conditions on the realization of null, pronominal, or verbal copulas in Arabic, see Fassi Fehri (1993) in particular.

**11.** Cf. Kayne 2000, Ch. 2 (originally published in 1985).

**12.** On the analysis of "have" as be+P/D, see Kayne (2000, Ch. 7, originally published in 1993), relying on previous work by Szabolcsi, Freeze, and Benveniste (see the references cited there).

**13.** In fact, Participle Agr may coincide with AgrO, as in (11), but not necessarily. Kayne was originally dealing with Participle Agr, but Chomsky (1995) generalizes AgrO to include it. See Siloni (1997) for evidence that AgrO is absent in Hebrew participles. As for the uniformity of the Agr chain with "be", but not "have", a similar conclusion has been reached by Iatridou, Anagnostopoulou, and Izvorski (2001) for e.g. Bulgarian, based on a personal observation made to them by Peter Svenonius. For these reasons, preference will be given to the use of (finite) Agr1 vs. non-finite (participle) Agr2, instead of AgrS and AgrO, although they are sometimes used interchangeably.

**14.** Interestingly enough, Marantz (p.c.) observes that these are the essential readings that obtain with thematic verbs in the Present in English (but not with the copula).

**15.** See in particular Fassi Fehri (1993) for detail.

**16.** These splits obtain via fissioning (or delinking) mechanisms in syntax, along the lines proposed in Fassi Fehri (1996b, 2000).

**17.** Morris Halle (p.c) has informed me that the analysis of *-er-* as auxiliary "be" is disputable. For discussion, see Giorgi & Pianesi (1997).

**18.** This example is taken from Embick (2000a), but the glosses are mine. As the author explains, Pass arises below T/Agr, but must readjust to be placed in T/Agr. If Imp Asp is realized as zero, Past realized by *-baa*, and the ending realizes a mixture of Pass and Agr, then the glosses give an approximate segmentation.

**19.** Note that the fact that Agr2/T2 is adjectival in GP and that it dominates Voice does not question its internal aspectuo-temporal nature, just as the dominance of finite temporal Voice by Agr in the complex Arabic tenses does not question its temporality. I return below to the role played by Height in determining language variation in Voice.

**20.** The same is true of the present passive, which I omit here. The examples are taken from Rivero (1990), except when signaled otherwise.

**21.** This example is taken from Joseph and Philippaki-Warburton (1987).

**22.** Rivero does not explain however why Pass does not stay as an affix on the thematic V, like in MG. It seems to me that the absence of this option is dictated by the necessity for Pass to interact with Asp. Similar interactions are found in Arabic, which led me to label the projection Vasp (for Voice and Asp) in Fassi Fehri (1988).

**23.** Benmamoun (2000) mistakenly analyzes this prefix as Asp progressive, despite evidence to the contrary. For an analysis close to mine, see Brustad (2000).

**24.** The bare prefixed form can also be used in embedded contexts, with an anaphoric Realis Mood.

**25.** See e.g. Heine (1993). My purpose here is to give appropriate motivation for the type of FC I am dealing with, as well as the structural mechanisms available for analysis. Cognitive factors may be playing a role in changes, but it is not clear how a cognitive approach would predict the right formal available changes in Ls, to the exclusion of others.

**26.** If the Present Perfect is analyzed on a par with other perfects as temporal complex, as I will assume, then Pass Perfect is of 3-level complexity. I disregard here the possibility of taking the source of analysis in Pass Perfect to be, for example, the avoidance of Agr mismatches, because this does not apply to Pass Imperfect.

**27.** In fact, Ackema argues that Pass in AG is higher than I; see his structures (135) and (137) for simple and perfect passives respectively (149-150). I leave the question of ordering Pass and I aside since it does not bear on the question I am addressing here.

**28.** See the references cited in Ackema (ibid).

**29.** Thanks to Ken Hale and John Lumsden for these judgments.

**30.** Another issue was whether the derivation is lexically or syntactically driven, a distinction that Borer (1998), among others, have argued to be dispensible.

**31.** The essential facts are the following:

i. In English, eventive and stative passives both surface as participles (with identical suffixes).

ii. In Chichewa, both eventive and stative can be just as verbal as anything else.

iii. In Creek, stative is as verbal as anything else.

See also Embick (2000a) for similar arguments based on Latin.

**32.** He also argues for the necessity of using Selection for appropriate characterization, a question that I will not address here.

**33.** This distinction appears to be clear-cut in the case of passive adjectives, but not in the case of active adjectives. For example, (58) embedded under *kaana* is solely interpreted as a stative complement of the copula, whereas (57) may have a coumpound tense interpretation, similar to that in (60). But in Modern Standard Arabic, the dominant tendency is to use finite verbs for complex Ts.

## References


Ackema, P. (1999). *Issues in Morphosyntax*. Amsterdam: John Benjamins.

Baker, M., Johnson, K. & Roberts, I. (1989). Passive arguments raised. *Linguistic Inquiry, 20.2*.

Beard, R. (1995). *Lexeme-Morpheme based Morphology.* Albany, NY: Suny Press.

Beard, R. (2001). How morphological asymmetry predicts the fusion-agglutination continuum. Paper presented at the *Asymmetry Conference 2001*. University of Quebec of Montréal.

Beedman, C. (1982). *The Passive Aspect in English, German, and Russian*. Tübingen: Gunter Narr Verlag.

Benmamoun, E. (2000). *The Feature Structure of Functional Categories.* Oxford and NY.: Oxford Univ. Press.

Borer, H. (1998). Deriving passive without theta roles. In S. G. Lapointe et al. (Eds.), *Morphology and Its Relation to Phonology and Syntax.* Stanford: CSLI Publications.

Bresnan, J. (1982). The passive in lexical theory. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*. Cambridge Mass: MIT Press.

Brustad, K. E. (2000). *The Syntax of Spoken Arabic.* Washington D.C.: Georgetown University Press.

Bussman, H. (1996). *Routledge Dictionary of Language and Linguistics.* London: Routledge.

Chomsky, N. (1995). *The Minimalist Program.* Cambridge Mass: MIT Press.

Cinque G. (1999). *Adverbs and Functional Grammar.* Oxford and N.Y.: Oxford University Press.

Comrie, B. (1976). *Aspect.* Cambridge: Cambridge University Press.

Comrie, B. (1985). *Language Universals and Linguistic Typology: Syntax and Morphology*. 2nd ed. N.Y.: Basic Blackwell.

Di Sciullo, A. M. (1999). The local asymmetry connection. *MITWPL, 35*. Cambridge, Mass: MIT Press.

Embick D. (2000a). Features, syntax, and categories in the Latin Perfect. *Linguistic Inquiry, 31.2*.

Embick D. (2000b). Participial structures and participial asymmetries. Ms., Cambridge, Mass: MIT Press.

Enç, M. (1991). On the absence of the present tense morpheme in English. Ms., Madison: University of Wisconsin.

Fassi Fehri, A. (1988). Arabic passive affixes as aspectual predicates. Ms., Cambridge: MIT and Rabat: Faculty of Letters.

Fassi Fehri, A. (1993). *Issues in the Structure of Arabic Clauses and Words.* Dordrecht: Kluwer Academic Publishers.

Fassi Fehri, A. (1996a). Nominality, VS/SV alternations and generalized checking. *Linguistic Research, 1.1*. Rabat: IERA.

Fassi Fehri, A. (1996b, 2000). Distributing features and affixes in Arabic subject verb agreement paradigms. *Linguistic Research, 1.2*. Also In J. Lecarme, J. Lowenstamm, and U. Shlonsky (Eds.), *Research in Afroasiatic Grammar*. Amsterdam: John Benjamins.

Fassi Fehri, A. (2000). Temporal variation across heights. Paper presented at the *Paris Roundtable on Tense and Aspect*. To appear in the Proceedings.

Giorgi, A. & Pianesi, F. (1991). Towards a syntax of temporal representations. *Probus, 3.2*.
Giorgi, A. & Pianesi, F. (1997). *Tense and Aspect.* Oxford and N.Y.: Oxford Univ. Press.
Guéron, J. (1995). Chaînes temporelles simples et structures auxiliaires. In P. Boucher et al. (Eds.), *Rencontres: Études de syntaxe et de morphologie*. Paris: Univ. of Paris X-Nanterre.
Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (Eds.), *The View from Building 20*. Cambridge, Mass: MIT Press.
Heine, B. (1993). *Auxiliaries: Cognitive Forces and Grammaticalization.* Oxford and N.Y.: Oxford Univ. Press.
Iatridou S., Anagnostopoulo, E. & Izvorski, R. (2001). Observations about the form and meaning of the perfect. In M. Kentowicz (Ed.), *Ken Hale: A Life in Language*. Cambridge, Mass: MIT Press.
Joseph, B & Philippaki-Warburton, I. (1987). *Modern Greek.* London: Croom Helm.
Kayne, R. (2000). *Parameters and Universals.* Oxford and N.Y.: Oxford Univ. Press.
Kratzer A. (1996). Severing the external argument from its verb. In J. Rooryck and L. Zaring (Eds.), *Phrase Structure and the Lexicon*. Dordrecht: Kluwer Academic Publishers.
Levin, B. & Rapaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry, 17.4*.
Marantz, A. (1997). No escape from syntax. In *Upenn WPL.* University of Pennsylvania.
Marantz, A. (2000a). Reconstructing the lexicon domain with a single generative engine. Ms., Cambridge, Mass: MIT Press.
Marantz, A. (2000b). The universality of root and pattern morphology. Ms., Cambridge, Mass: MIT Press.
Matthews, P.H. (1972). *Inflectional Morphology.* Cambridge: Cambridge University Press.
Ouhalla, J. (1991). *Functional Categories and Parametric Variation.* London: Routledge.
Rivero, M. L. (1990). The location of nonactive voice in Albanian and Modern Greek. *Linguistic Inquiry, 21.1*.
Siloni T. (1977). *Noun Phrases and Nominalizations*. Dordrecht: Kluwer Academic Publishers.
Tsimpli, I. M. (1989). On the properties of the passive affix in Modern Greek. *UCL Working Papers in Linguistics, 1*.
Wasow, T. (1977). Transformations and the lexicon. In P. W. Culicover et al. (Eds.), *Formal Syntax*. NY: Academic Press.

# Asymmetry and linearization in phonology

Eric Raimy
Haverford College / Swarthmore College

This paper investigates whether the *asymmetry hypothesis* proposed by Kayne (1994) is relevant to phonology. A broad interpretation of the asymmetry hypothesis is that *linear precedence* is derivable from and dependent on *asymmetrical* relations. Any connection between precedence in linear ordering and asymmetry in representations has important ramifications for phonological theory because the types of phonological representations that are considered are shaped by these constraints. Barring a brief discussion in Cinque (1996) there has been no investigation of what the asymmetry hypothesis can tell us about linear ordering in phonology or the nature of phonology in general. The main focus of this paper is to discus the relevance of the asymmetry hypothesis for the representation of linear order in phonology. Specifically, the issue of when *non-asymmetrical* precedence representations occur and how they are dealt with in the phonology in light of proposals made in Raimy (2000b) is investigated. Following from this discussion a comparison of how non-asymmetrical representations are handled in syntax and phonology is outlined. This path of investigation allows both similarities and differences between representations in syntax and phonology to be identified and this contributes to our better understanding of universal grammar.

## 1. Asymmetry in phonology

Traditional assumptions on the relationship of precedence in phonological representations indicate that it has the following characteristics.

*Precedence is non-transitive.* Precedence in phonological representations is neither necessarily *transitive* or *atransitive* because of the nature of adjacency in phonology. Long distance effects in phonology as evidenced by vowel harmony across consonants indicate that some processes require precedence to be *transitive*. This type of precedence requires that if A precedes B and B precedes C, then A precedes C to be true. Processes that require strict adjacency between segments indicate that precedence in phonology may also be *atransitive* in that if A precedes B and B precedes C, then A does not precede C must be true. Because precedence appears to be transitive in some cases and atransitive in others,[1] it is properly characterized as *non-transitive*, meaning that strict transitivity is not required.

*Precedence is irreflexive* in that if A precedes B, then A is not B must be true. This characteristic is implicitly assumed since the idea of reflexivity of precedence has not been pursued in phonological representation. One possible interpretation of reflexivity of precedence in phonology is the representation of geminates as a single melody linked to two timing slots. In a sense, due to the multi-tiered representation of geminates a unit of melody does 'precede itself' when the timing tier is considered. This appearance of reflexiveness of precedence results from autosegmental units being allowed to be associated with more than a single unit on another tier. Within a single tier in an autosegmental representation, precedence can only be irreflexive.

*Precedence is asymmetrical* meaning that if A precedes B, then B does not precede A must be true. Simply put, there is no language where [kæt] = [tæk]. This is the core attribute of phonological representations that allows the relevance of Kayne's (1994) proposals to be seen in phonology. The asymmetry hypothesis holds in phonology because all precedence structures are ultimately asymmetrical.

The characteristics of precedence discussed above are presently understood as static and inviolable constraints on the well-formedness of a phonological representation. A phonological representation may start out as ill-formed due to the presence of a floating feature or if the position that morphemes are not linearly ordered in the input is assumed. These situations do not violate any of the proposed characteristics of precedence because both of them arise from two pieces of phonological representation not having any relation of precedence between them. Since there is no precedence relation, the conditions placed on precedence relationships can not be violated. We can thus conclude that all precedence relationships respect all of the constraints on phonological representations derived from these characteristics. This situation is similar to Kayne's (1994) interpretation of the LCA where all syntactic representations are required to be well-formed with respect to asymmetry. While this static view of asymmetrical properties of phonological representations is the prevailing model, it is not necessarily the only model of phonological representations.

## 2. Non-asymmetrical precedence structures in phonology

Raimy (2000b) argues that precedence relationships must be explicitly notated in all phonological representations. This point causes a change in the way phonological representations are presented as seen in (1). '→' indicates the relationship of *precedence* in the representations proposed by Raimy (2000b). Note that '#' and '%' represent the null set so they should be read as 'nothing precedes X' = # and 'X precedes nothing' = %. See Raimy (2000b) for a discussion for the need of both of these units to define the beginning and end of a word.

(1) a. kæt
    b. # → k → æ→ t → %

The representations in (1) are actually equivalent in the information about the precedence structure of the word in question if the characteristics of precedence that are discussed in Section 1 are assumed. The difference between (1a) and (1b) is only in whether the precedence information is explicitly presented. Given the strong constraints on precedence relations in the static view presented in Section 1, there is no pressing reason to complicate our graphic presentation of phonological representations because the only way (1a) can be interpreted is as (1b).

One effect of the lack of explicit precedence notation in phonological representations is that the assumptions about the characteristics of precedence discussed in Section 1 have not been questioned. The traditional phonological representation in (1a) limits us to only considering phonological representations that follow the constraints on precedence outlined in Section 1 because if these assumptions are abandoned, the precedence structure in (1a) becomes very difficult to interpret. If explicit precedence notation is added back into phonological representations as in (1b), we can question whether the characteristics of precedence that have been assumed are correct or not.

Given explicit precedence notation, we can consider phonological representations that violate the conditions on precedence in Section 1. Raimy (2000a, b) proposes that reduplication is best represented by precedence structures that are 'looping' in nature. Consider the representation in (2).

(2) # → k → æ → t → % ‘cat-cat’

The representation in (2) violates the conditions of *asymmetry* and *irreflexivity* for precedence if we assume the *transitive* version of precedence. If precedence is a transitive relation, then in (2) every segment precedes itself and this violates the irreflexive characteristic of precedence. It also follows that the entire precedence structure is *symmetrical* because every segment precedes every other segment since /k/ precedes /æ/ and /æ/ precedes /k/ through transitivity, etc.

Within a model of phonology that allows representations like (2), non-surface[2] representations are not necessarily interpretable at the relevant interface. One aspect of phonology is to ensure that representations stored in memory are converted to representations that are interpretable at the phonetics interface. (2) is not interpretable at the phonetics interface because the overall precedence ordering in the representation is not *asymmetrical*. This is one point where phonology is indebted to Kayne (1994) which proposes the connection between asymmetry and linear order.

Analogous to the syntactic LCA (Kayne 1994), phonology contains a *linearization* process which ensures that representations have *asymmetrical* precedence structures and are thus interpretable at the phonetics interface. Linearization in phonology repeats segments within a ‘loop’ in order to eliminate asymmetrical precedence relations. Consider the representations in (3).

(3) a. # → k → æ → t → %

b. # → k → æ → t → k → æ → t → %

Both (3a) and (3b) contain the same *types* of precedence relations. The term *type* here captures the fact that although (3a) and (3b) are different along some dimensions (number of segments and distinct environments that each segment appears in) they are also the same if we only refer to unique precedence relations. The specific types of unique precedence relations found in (3ab) are presented in (4).

(4) {# → k} {k → æ} {æ→ t} {t → %}{t → k}

The difference between (3a) and (3b) is to be found in the *tokens* of precedence relations in each representation. The notion of tokens of precedence relations is directly tied to the number of segments used to represent the precedence information. (3a) and (3b) have different numbers of segments which results in a different number of tokens of precedence relations. Consider the lists of tokens of precedence relations in (5).

(5) a. tokens for (3a) [# → k] [k → æ] [æ→ t] [t → %][t → k]
b. tokens for (3b) [# → k] [k → æ] [æ→ t] [t → k'] [k' → æ'] [æ' → t'] [t' → %]

By repeating segments, which creates new tokens of precedence relations, linearization allows all of the precedence information in (3a) to be contained in (3b) without containing any non-asymmetrical precedence relations. Crucial to understanding linearization in phonology is that copies (i.e. different tokens) of particular segments (same type) created by linearization do not have a relationship of *identity*[3] or any other type of dependency relation between them. The precedence structure in (5b) and the graph in (3b) do not contain any symmetrical precedence relations because each segment is a distinct entity. The fact that a /t/ precedes a /k/ and the /k/ precedes another /t/ does not instantiate a symmetrical precedence relation because the two /t/s (and two /k/s) in question are not the same entity. They are distinct segments and have distinct precedence environments which do not create a non-asymmetrical precedence structure.

Linearization in phonology creates and repeats distinct tokens of types of precedence relations in order to eliminate non-asymmetrical precedence structures. The fact that phonology chooses to repeat segments in order to remedy non-asymmetrical representations (as opposed to deleting segments) is an important point to note, especially when the way that non-asymmetrical structures in syntax are remedied (discussed in Section 4).

To summarize this section, four main points have been made. The first is that restrictions on possible precedence structures in phonology do not hold at all levels of representation. Phonological representations stored in memory or assembled in the morphology may violate some condition on what is interpretable at the phonetics interface. This leads to the second point which is that a linearization process in the phonology ensures that representations are interpretable with respect to precedence structures at the phonetics-phonology interface. Converting abstract memorized forms into representations that can be utilized by the phonetics module and consequently turned into motor control programs is one of the fundamental aspects of phonology. The behavior of the linearization process is the third main point of this section. Linearization eliminates non-asymmetric precedence structures through repetition of segments which preserves the overall organization of a precedence structure while not causing problems of interpretation for the phonetics module. The final point is that we can see concrete effects of symmetrical precedence structures and linearization in reduplicated forms. The repetition of segmental material that we recognize as reduplication is the result of the linearization of a non-asymmetric precedence structure which was created in the morphological component.[4]

## 3. Linearization as an economy-based process

One of the main points of the previous section is that there is a linearization process in the phonology which remedies non-asymmetrical precedence structures through repetition. This section investigates the characteristics of the linearization process in further detail and argues that linearization is an economy-based process. Two main principles of economy that are relevant for linearization are presented in (6).

(6) a. Shortest Path- produces the precedence graph with the least number of tokens of precedence relations from the input
b. Completeness- produces a precedence graph that uses as many of the types of precedence relations from the input as possible

The interaction between the principles in (6) creates the optimization characteristic of linearization in that structures that are created by linearization employ only as many tokens of precedence relations that are needed to use all the types of precedence relations in the input.

Linearization has characteristics in common with the *containment model* of Prince and Smolensky (1993). Possible linearizations are strictly constrained by the input precedence structure because only precedence relationships that are present in the input structure can be present in the output structure. No novel precedence types are ever added or created by the linearization process itself. Another basic well-formedness condition on the output of linearization is that every precedence structure will have a beginning ([# → X]) and an end ([Y → %]). If a precedence structure does not have a beginning or an end, it is not interpretable at the phonetics-phonology interface.

Given this initial sketch of the properties of the linearization process, consider what occurs when a completely asymmetrical linear precedence structure is linearized as presented in (7).

(7) *input*: a. # → k → æ→ t → %
*outputs*: b. # → k → æ→ t → %
c. **# → t → æ→ k → %
d. **# → k → æ
e. **# → d → a → g → %
etc.

The input in (7a) will be linearized as (7b) and both the Completeness (in the output there is at least one token of each type of precedence relation from the input) and the ShortestPath (there is only a single token of each type of precedence relation from the input present in the output) are maximally satisfied. Other possible linearizations that could be considered by linguists are not entertained by the linearization process because they contain precedence relationships that are not present in the input structure. (7c) is the mirror structure of the input and actually contains none of the precedence relations from the input. Because of this, it can not be produced by the linearization process as a legitimate linearization. Structures that are not legitimate linearizations of the input structure are marked by double asterisks (**). (7d) is also illegitimate specifically because it does not have an 'end' (there is no [t → %] which is the only legitimate ending according to the input) in the output precedence graph. (7e) and other analogous structures (e.g. 'plane', 'transcendentalism', 'cats', etc.) are also illegitimate because they contain segments that are not present in the input structure which automatically causes there to be precedence relations present in the output that do not occur in the input.

Precedence structures that are not asymmetric will have non-trivial linearizations because these types of structures present situations where ShortestPath and Completeness may be in conflict. Due to this conflict, there will be more than a single possible way to linearize the input. Consider the linearization possibilities for a looping structure given in (8). Note that illegitimate linearizations are excluded from the discussion here since there are actually multiple possible linearizations that compete for the best satisfaction of both principles of economy.

(8) input: a. # → k → æ → t → %

output: b. *# → k → æ → t → %

c. # → k → æ → t → k → æ → t → %

d. *# → k → æ → t → k → æ → t → k → æ → t → %

etc.

The multiple possible linearizations presented as outputs in (8b-d) highlight the conflict in the economy principles. (8b) is the smallest possible output precedence structure that can be created by the linearization process. (8b) has a beginning and an end but the principle of Completeness is not satisfied because there is no token of the [t → k] precedence relation (the 'loop back') in the output. The omission of this particular precedence relation is why

(8b) does not show reduplication. (8c) is the most economic output that linearization can produce. This representation is not the smallest possible precedence structure because it has seven ([# → k] [k → æ] [æ→ t] [t → k'] [k' → æ'] [æ' → t'] and [t' → %]) tokens of precedence relations while (8b) only has four but it is the smallest representation that satisfies Completeness. There is at least one token of each type of precedence relation from (8a) in (8c). At this point, we can see that the economy principle of Completeness overrides ShortestPath when there is a conflict between them. (8d) and all other possible output structures that use the [t → k] precedence link more than once equally satisfy Completeness but are larger precedence structures than (8c). Thus, these gratuitously large precedence structures are not the most economical possible linearization.

Indonesian provides an excellent case study which further illuminates the economy principles of linearization. Consider the data in (9) (taken from Uhrbach 1987; McCarthy & Prince 1995).

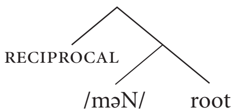
| | | | | |
|---|---|---|---|---|
| (9) | a. | potoŋ | mə-motoŋ-motoŋ | 'to cut (intensive/repetitive)' |
| | | tulis | mə-nulis-nulis | 'to write (intensive/repetitive)' |
| | | duduk | mən-duduk-duduk | 'to occupy ('sit' intensive/repetitive)' |
| | | isi | məŋ-isi-isi | 'to fill with various things' |
| | b. | pukul | pukul-məm-ukul | 'to hit (reciprocal)' |
| | | tari | tari-mən-ari | 'to dance (reciprocal)' |
| | | hormat | hormat-məŋ-hormat-i | 'to respect (reciprocal)' |
| | | ambil | ambil-məŋ-ambil | '(give and) take' |

The relevant aspects of the data in (9) are as follows.[5] Total reduplication is present as the result of both 'intensive/repetitive' (9a) and 'reciprocal' morphology (9b). The prefix /məN/ is not reduplicated as part of the total reduplication in either (9a) or (9b) The position of the prefix /məN/ alternates between prefixing to the base and reduplicant in (9a) and being interposed between the base and reduplicant in (9b). How to account for the interposing pattern of reduplication in (9b) is the important question in this data.

The difference in the placement of /məN/ in (9) and an explanation of how /məN/ becomes interposed between base and reduplicant is derivable from the difference in morphological structure of (9a) and (9b) and the principles of economy. The morphological structure of the forms in (9) is basically a question of which morpheme has the widest scope. The morphological structure for the prefixing pattern in (9a) is found in (10a) where /məN/ has the widest scope, including both the 'intensive/repetitive' morpheme and the root within it. The interposing pattern in (9b) has the morphological structure in (10b) where the ‘reciprocal’ morpheme has scope over both /məN/ and the root.

(10) a.


b.




Given the morphological structures in (10) and the assumption of cyclic spell-out of the morphosyntactic features, we arrive at the derivations of the phonological representations (precedence graphs) for (9a) and (9b) in (11). (See Raimy 2000b for a full discussion of the concatenation of morphemes given the precedence structures under discussion.)

(11) a. PREFIXED

i. *root* # → d → u → d → u → k → %

ii. *INTENS* # → d → u → d → u → k → %

iii. */məN/* m → ə → n
↑
# → d → u → d → u → k → %

b. INTERPOSED

i. *root* # → a → m → b → i → l → %

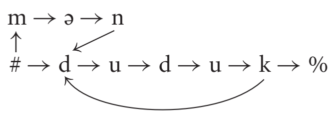
ii. */məN/* # → a → m → b → i → l → %
↓
m → ə → ŋ

iii. *RECIP* # → a → m → b → i → l → %
↓
m → ə → ŋ

Both INTENSIVE and RECIPROCAL morphology trigger total reduplication which adds a precedence link from the last segment of the phonological string to the first segment. This creates a loop from the *end* to the *beginning* of the form. Depending on whether /məN/ has already been spelled-out, /məN/ may be contained within the loop or not. As (11) shows, fundamentally different precedence structures are created for forms with RECIPROCAL and INTENSIVE morphology in Indonesian even though they are constructed from the same pieces. The different precedence structures result from the different order in which the same pieces are put together.

In (11a), the intensive morpheme is spelled-out when only the root has been spelled-out. This produces the precedence structure in (11aii) where /məN/ is concatenated to a representation that already contains a loop. Consequently, in (11aiii) /məN/ is not contained within the loop.

(11b) shows a different order of spell-out/concatenation of morphemes. In this form, /məN/ is spelled out first and concatenated directly to the root as shown in (11bii). After this, the reciprocal morpheme is spelled-out which adds a precedence relation from the last segment of the representation to the first segment. Since /məN/ has been concatenated to the root, /m/ is now the first segment thus causing /məN/ to be included inside the loop as shown in (11b-iii). This is a different precedence structure from (11a-iii) and we should expect different surface linearizations to occur for these two representations. Linearization of the structure in (11a-iii) (the prefixed /məN/ pattern) follows directly from the previous discussion of interaction of the economy principles. (12) presents the occurring and some possible alternative linearizations.

(12) a.



b. mən-duduk-duduk
c. *duduk-duduk
d. *mən-duduk
e. *duduk
f. *mən-duduk-duduk-duduk

(12b) is the occurring surface form and it is the smallest precedence structure possible (14 precedence relations) that uses both the prefix /məN/ and the 'loop'. Completeness is maximally but not completely satisfied by this output because there is no [# → d] precedence relation in the output. There is no way that both the /məN/ affix and this precedence link can be used given the structure in (11aiii). Linearization must make a choice between using the [# → m] link and the using [# → d] link. Choosing to use the [# → d] link produces (12c) a representation in which Shortest Path is minimally violated but there are more violations of Completeness than (12b). Other possible linearizations, as in (12d-e), also produce smaller precedence structures than (12b) but utilize less of the precedence structure thus violating Completeness more than (12b). This rules them out from being the most economical linearization. (12f) satisfies Completeness to the same level as (12b) but is a larger precedence structure. This fact makes (12f) less economical with respect to ShortestPath.

The precedence graph for (11biii) (the interposed /məN/ pattern) presents a more complicated interaction between the economy principles in linearization. Although the interaction is more complicated, the surface form for (11b-iii) is generated by the present principles of economy.

(13) a.

$$
\begin{tikzcd}[column sep=small]
\# \arrow[r] \arrow[d] & a \arrow[r] & m \arrow[r] & b \arrow[r] & i \arrow[r] & l \arrow[r] \arrow[dlllll, bend left] & \% \\
m \arrow[r] & \text{ə} \arrow[r] & \text{ŋ} \arrow[ul]
\end{tikzcd}
$$

b. ambil-məŋ-ambil
c. *məŋ-ambil-məŋ-ambil
d. *ambil
e. *məŋ-ambil

Given the precedence structure in (13a), (13b) is the smallest precedence structure in which Completeness is maximally satisfied. As discussed in the linearization of (11aiii), there can only be one 'word initial' segment so there is a choice in between using the [# → a] and using [# → m] precedence link. There is no way to use both of these links given the *containment* nature of the linearization process. (13b) is produced by following the [# → a] link as a 'short cut' to using the loop back link. Because /məN/ is within the looping back link, the short cut through [# → a] is not detrimental to the overall satisfaction of Completeness because there will always be at least one unused precedence link.

(13c) shows what occurs when the [# → m] is utilized by the linearization process. This path causes /məN/ to be present in the output twice leading to a larger precedence structure than (13b) with no gain in Completeness. Because both (13b) and (13c) leave one precedence link unused this makes both linearizations equal with respect to satisfying the Completeness economy principle. Since Completeness does not distinguish between the two linearizations, ShortestPath determines that (13b) is the most economic linearization since it is smaller than (13c).

Other possible linearizations such as (13d–e) do not maximally satisfy Completeness and are thus ruled out. Linearizations that are larger than (13c) (**məŋ-ambil-məŋ-ambil-məŋ-ambil*, **ambil-məŋ-ambil-məŋ-ambil*, etc.) do not improve satisfaction of Completeness so the only result produced by these forms is more violation of ShortestPath.

We can now see that the curious pattern of interfixing the /məN/ prefix in Indonesian is derived from the general language universal principles of linearization and morphological scope/spell-out in the morphology. This investigation into the principles of linearization indicate that it is an optimization process that occurs in the phonology. Linearization is driven by two principles of economy,[6] Shortest Path and Completeness, and Completeness overrides Shortest Path when these economy principles are in conflict.

## 4. Linearization in syntax and phonology: Why is phonology different?

Both syntax and phonology have the characteristic that asymmetrical representations are required at the relevant interface. Syntax and phonology also share the characteristic that non-asymmetrical representations can be built or occur during a derivation. Section 2 of this paper discusses how the phonology contains non-asymmetrical looping precedence structures that were built within the morphology component. Syntax builds non-asymmetrical structures when movement occurs (the copy theory of movement, Chomsky 1993). To see this fact about syntax, consider the example in (14) (taken from Nunes 1999).

(14) a. [$_{TP}$ John$^{i}$ [$_{T'}$ was + T [$_{VP}$ kissed John$^{i}$]]]
b. *John was kissed John
c. John was kissed

Copy and movement of 'John' from the VP in (14a) creates symmetrical and reflexive c-command relations according to Nunes (1999: 225) because this movement causes 'John' to c-command 'kissed' but 'kissed' c-commands the lower copy of 'John'. This is a case of symmetrical c-command which violates the LCA. In addition to this, a reflexive c-command relationship is also created because 'John' in subject position c-commands 'John' in object position. Nunes (1999) argues that the general response to chains that create symmetrical and reflexive c-command structures in violation of the LCA (Kayne 1994) is to delete part of the chain at PF.[7]

At this point, we have reached a paradox if asymmetry and linearization is to be a cross-modular factor in UG. Specifically, syntax deletes elements in order to change symmetrical representations into asymmetrical ones and phonology repeats elements to achieve a similar conversion. One possible response to this paradox is to simply dismiss it as another basic difference between the syntactic and phonological modules. A deeper explanation of this paradox can be formulated if we consider how the notion of *identity* operates over representations in each module.

Syntax requires a relationship of identity in order to allow chains to be formed. Chains instantiate symmetrical c-command relations because the elements of a chain are identical from the perspective of the syntax. This situation causes the moved category to c-command itself, and for intervening categories, to c-command and be c-commanded by the moved category. Consider the example in (15) which illustrates this point.

(15) a. John$^i$ wants John$^j$ to go
b. *John$^i$ wants John$^i$ to go

The difference in grammaticality between (15a) and (15b) can be explained by considering whether the second 'John' in each sentence in (15) refers to the same person as the subject or not. (15a) is only acceptable when the two 'Johns' refer to different individuals and thus do not form a chain. Since these two instances of 'John' are not *identical*, no symmetrical or reflexive c-command relations exist in this sentence. (15b) is ungrammatical because the two instances of 'John' are identical forming a chain which creates a symmetrical c-command relation which violates the LCA.

An explanation of the paradox of syntax deleting parts of a representation to achieve asymmetry and phonology repeating parts to obtain the same goal can be developed based on the role that *identity* plays within each grammatical module. Simply, phonology behaves differently from syntax because there is no dependency between copies of segments. This lack of dependency means that there is no sense of identity between *tokens of segments* produced by linearization.[8] The claim here is that phonology does not require a function of *token identity* that relates two distinct segments (prosodic units, features, etc.) and requires them to be (or evaluates whether they are) identical along some dimension. This type of *token identity* is the core of the correspondence model (McCarthy & Prince 1995) of reduplication which formalizes Wilbur's (1973) Identity Constraint. The main liability of this type of model of reduplication is that since *identity* is an arbitrary relation that can hold between any two segments whether they have any similarity of features at all it produces a too powerful model of phonology which causes problems for acquisition. See Idsardi and Raimy (forthcoming) for discussion of this theme.

The model of reduplication argued for in Raimy (2000a, b) accounts for identity effects in reduplication without resorting to *token identity* by claiming that identity effects in reduplication (overapplication and underapplication, including backcopying) only occur when a rule applies to a phonological representation that contains a 'loop'. The loop is crucial to this explanation because it presents a situation where a segment can appear in two distinct environments at the same time. When these multi-environments diverge along some dimension, identity effects are produced because after linearization has occurred, each copy of a segment may appear in a distinct environment.

To show how these multi-environments adequately capture identity effects in reduplication, consider the behavior of nasal spread in Malay (Kenstowicz 1981) presented in (16).

(16)

a. aŋẽn ‘wind’ ãŋẽn- ãŋẽn ‘unconfirmed news’

b. # → a → ŋ → e → n → % ⇒ # → ã → ŋ → ẽ → n → % *nasal spread*

c. # → ã → ŋ → ẽ → n → ã→ ŋ → ẽ → n → % *linearization*

McCarthy and Prince (1995) identify this pattern of rule application in reduplication as one that shows backcopying. Backcopying occurs in this pattern because the word initial vowel in the reduplicated form becomes nasalized even though it is never preceded by a nasal segment. In other words, there does not appear to be any way to cause the nasalization of the word initial vowel without resorting to some interpretation of *identity* between this vowel and the corresponding word internal vowel which occurs in a nasalizing environment.

Raimy (2000a) argues that the backcopying identity effect results from the nasal spread rule in Malay applying to the pre-linearization representation in (16b). The representation in (16b) presents a case of a multi-environment situation since /a/ is ‘at the beginning of the word’ and ‘following /n/’ at the same time. Symmetrical precedence structures create multi-environments which lead to opaque identity effects after linearization. The surface form in (16c) is derived, according to Raimy (2000a), by allowing the nasal spread rule to apply to the prelinearized representation in (16b) even though only one part of the multi-environment triggers nasalization. Linearization then places one copy of the now nasalized vowel in word initial position, (16c), and the backcopying effect is accounted for without resorting to any kind of *identity constraint* or *transderivational device*. Identity effects in reduplication are cases of opacity in the Raimy (2000a, b) model.

Non-identity effects will occur in reduplicated forms when a rule applies after linearization has already converted a loop into a repeated region of segments. Consider the example of coda devoicing in Washo (Jacobson 1964 as cited in Kager 1999: 231). Syllable boundaries are marked with ‘ . ’ in (17a) and with parentheses in (17c, d).

(17) a. /RED+wed-i/ wet.we.di 'it's quacking'

b. #→ w → e → d → %

i

c. # (w → e → d) → (w → e) → (d → i) → % *linearization*

d. # (w → e → t) → (w → e) → (d → i) → % *coda devoicing*

All that needs to be done, in order to account for this pattern of rule and reduplication interaction, is to allow linearization to occur before coda devoicing applies. This ordering produces the representation in (17c) for coda devoicing to operate on. Since there are now two distinct copies of the /d/ in (17c), coda devoicing can operate on them independently. Application of this rule causes devoicing of the first /d/ because it occurs in a coda while the second /d/ is unaffected because it occurs in an onset. The two copies of underlying /d/ diverge in quality in (17d) as a result of coda devoicing because there is no *identity* between copies of segments in phonology.

The importance of non-asymmetrical relations in both syntax and phonology can now be recognized and a fundamental difference between syntax and phonology has been identified from this point. The syntactic component eliminates non-asymmetrical relations among elements by deleting copies of the offending elements at PF. Phonology eliminates non-asymmetrical relations by copying and repeating elements during linearization. This difference between syntax and phonology can be derived from whether the grammatical module in question requires a function of *identity*. Syntax requires *identity* in order that chains be formed while phonology does not require any dependency of this sort. Consequently, a transderivational identity function is not required or justified by the necessity of accounting for reduplication. All identity effects in reduplication can be achieved through derivational ordering of processes and this results in a less powerful model of phonology.

## 5. Conclusion

This short paper has investigated the relevance of the *asymmetry hypothesis* proposed in Kayne (1994) to phonology. From this point of inquiry, we have identified some characteristics of *linearization* that appear to be part of universal grammar. These characteristics are the following: linearization ensures that representations at interfaces have strict asymmetrical linear ordering and linearization operates under principles of economy. Given these universal aspects of linearization, we have also identified module specific characteristics of linearization. Linearization will either delete or repeat parts of non-asymmetrical representations depending on whether the module has a function of identity or not. Syntax requires a function of identity so chains can be created and interpreted correctly, and this causes deletion to be the remedy for symmetrical representations. Phonology does not require a function of identity (contrary to McCarthy & Prince 1995) so repetition of segments in non-asymmetrical representations occurs. Finally, although linearization in both syntax and phonology operates under principles of economy, there is likely to be distinct principles that are relevant for each module. Because of the difference in representations in syntax and phonology (hierarchical trees of categories based on c-command vs. strings or precedence graphs of segments), it is unlikely that linearization in each module operates under the same principles.

## Notes

* I would like to thank Milan Rezac, Cedric Boeckx and the participants of the Asymmetry Conference for discussing the ideas here with me. All mistakes, misstatements or misunderstandings are my responsibility alone.

1. Whether a process requires a transitive interpretation of precedence or an atransitive interpretation appears to be specified as a parameter of a given process. See the discussion of this aspect of phonological processes in Odden (1994) for concrete examples of both long distance (i.e. transitive) and strictly local (i.e. atransitive) precedence requirements on phonological processes. If adjacency in phonology operates in a 'relativized' way along the lines suggested by Odden (1994), then precedence may be a purely *atransitive* relation and can be calculated on tiers of representation other than the timing (x-slot) tier which allows the presence of surface 'long distance' effects.

2. We can include input or underlying representations and any type of 'intermediate' (derivational level, sympathetic candidate (McCarthy 1999), etc.) representation in the phonology that is not actually sent to the phonetics interface. A bare output condition (Chomsky 1995) on phonological representations is that it must be a strictly linear representation that follows all of the restrictions on precedence in Section 1. In other words, precedence structures that are not asymmetrical are not interpretable by the phonetics interface.

3. I use the term identity here in the sense of Wilbur's (1973) *identity constraint* which states that two segments can be in a relationship of identity which can cause phonological processes to apply (or not apply) in an exceptional manner to ensure the two segments stay identical. This relation creates a dependency between two segments that in effect causes them to be the 'same' segment thus creating symmetrical precedence relations.

4. Not all reduplication must be created in the morphology component. Cases of 'inherent reduplication' (e.g. cases in Manam as argued for by Buckley 1999 or possibly cases of biliteral roots in Semitic as argued for by Gafos 1998) can be accounted for by simply allowing memorized phonological forms to contain loops.

5. The interaction of reduplication and nasal substitution seen in some of the forms in (9) will be ignored for the purposes of this paper. See Raimy (2000b: 99–112) for an analysis of this aspect of the Indonesian data.

6. A preference to use morphologically added material over underlying material may be an additional aspect of Completeness or an additional independent economy principle. See Raimy (2000b) for discussion.

7. Nunes (1999: 232–237) discusses cases where some intermediate traces are not deleted. A brief summary of Nunes' position is that these traces are not deleted because a morphological process has merged the trace with another constituent into a phonological word (following a suggestion in Chomsky 1995: 337). This process eliminates the trace from the LCA thus allowing it not to be deleted.

8. This kind of identity is distinct from *type identity* which may be required in phonology to evaluate certain types of processes as discussed by Reiss (this volume). The main difference between *type* vs. *token* identity is that *type identity* only evaluates whether two segments are the same with respect to a certain feature or set of features (i.e. Is this segment [coronal]?). *Token identity* requires an evaluation which determines whether two segments are copies of each other or share some other type of dependence between them (*correspondence* in McCarthy & Prince 1995 terms). Under this evaluation, *token identity* would allow two completely different segments to be *identical* (correspond in this example) without sharing a single feature. This situation is instantiated by some emergence of the unmarked analyses of reduplication where miscopying is preferred to faithful copying because a reduction of markedness is achieved (Alderete et al. 1999). *Type identity* does not allow this type of *identicalness* because what is being evaluated is whether two segments both have a feature or set of features (or possibly more complex conditions as argued by Reiss). So, if two segments do not share any feature in common, they can not be considered to be identical along any dimension.

## References

Alderete, J., Beckman, J., Benua, L., Gnanadesikan, A., McCarthy, J. & Urbanczyk, S. (1999). Reduplication with fixed segmentism. *Linguistic Inquiry, 30*, 327–364.

Buckley, E. (1999). Integrity and correspondence in Manam double reduplication. In P. N. Tamanji and K. Kusumoto (Eds.), *NELS 28, Vol. 1*, (pp. 59–68). Amherst: GLSA, University of Massachusetts.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale and S. J. Keyser (Eds.), *The View from Building 20: Essays in Honor of Sylvain Bromberger* (pp. 1–52). Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, G. (1996). The 'Antisymmetric' Programme: Theoretical and typological implications. *Journal of Linguistics, 32*, 447–464.

Gafos, D. (1998). Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory, 16*, 223–278.

Idsardi, W. & Raimy, E. (forthcoming). Reduplicative economy. In B. Vaux (Ed.), *Rules and Constraints in Contemporary Phonological Theory*. Oxford: Oxford University Press.

Jacobson, W. (1964). A grammar of the Washo language. Ph.D. dissertation, Department of Linguistics, University of California, Berkeley.

Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Kayne, Richard. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kenstowicz, M. (1981). Functional explanations in generative phonology. In D. L. Goyvaerts (Ed.), *Phonology in the 1980's*. Ghent: E. Story-Scientia.

McCarthy, J. (1999). Sympathy and phonological opacity. *Phonology, 16*, 331–399.

McCarthy, J. & Prince, A. (1995). Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory* (pp. 249–384). Amherst: GLSA, University of Massachusetts.

Nunes, J. (1999). Linearization of chains and phonetic realization of chain links. In S. D. Epstein and N. Hornstein (Eds.), *Working Minimalism* (pp. 217–249). Cambridge, MA: MIT Press.

Odden, D. (1994). Adjacency parameters in phonology. *Language, 70*, 289–330.

Prince, A. & Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms. Rutgers University, New Brunswick, and University of Colorado, Boulder.

Raimy, E. (2000a). Remarks on backcopying. *Linguistic Inquiry, 31*, 541–552.

Raimy, E. (2000b). *The Phonology and Morphology of Reduplication*. Berlin: Mouton de Gruyter.

Reiss, C. (2003). Towards a theory of fundamental phonological relations. In A.-M. Di Sciullo (Ed.), *Asymmetry in Grammar, Vol. 2: Morphology, Phonology, Acquisition*. Amsterdam/Philadelphia: John Benjamins.

Uhrbach, A. (1987). A formal analysis of reduplication and its interaction with phonological and morphological processes. Ph.D. dissertation, Department of Linguistics, University of Texas, Austin.

Wilbur, R. (1973). The phonology of reduplication. Ph.D. dissertation, Department of Linguistics, University of Illinois. Distributed by the Indiana University Linguistics Club, Bloomington, Indiana.

# Levels, constraints and heads

Harry van der Hulst and Nancy A. Ritter
University of Connecticut

## 1. Introduction

The organization and content of this article is as follows. In Section 2, we offer a general outline of a two-level (i.e. lexical, post-lexical), principles-and-parameters approach to phonology. Though derivational in the sense of having two levels,[1] the model is constraint-based at each level (for us constraints are 'principles', i.e., universal constraints, and 'parameters' are constraints that languages choose from a limited pool). Since the constraints are not extrinsically ranked, the model only uses 'hard' (i.e. unviolable) constraints. Having discussed the derivational side of our approach in Section 2, Section 3 presents our commitment to a theory of phonological representations that attributes a fundamental role to the asymmetrical relation of head-dependency. This section contains an example of the type of analysis that our approach makes available. In Section 4, we turn to a proposed refinement of parameter theory that, in fact, extends the role of the head-dependency asymmetry into the domain of parameter setting. While this proposal is primarily made in the spirit of developing our own approach to phonology, it can, at the same time, be seen as an answer to certain objections that have been raised against parameter theory by proponents of Optimality Theory (OT). Implicit in this latter remark is the point that we ourselves have not embraced the ideas of OT (especially in rejecting the notion of constraint ranking). In comparing our own constraint-based approach to OT, we will take the opportunity to examine and criticize the OT-device of language-specific (parochial) constraint ranking. Section 5 summarizes our main conclusions and indicates areas for further research.

## 2. The organization of phonology

Our goal in this section is to argue in favor of a two-level (lexical/post-lexical) constraint-based approach to phonology that incorporates the so-called principles-and-parameters framework. As we present our views, we argue against the use of parochial constraint ordering (as advocated in Optimality Theory, OT; Prince & Smolensky 1993).[2]

### 2.1 Phonological constraints

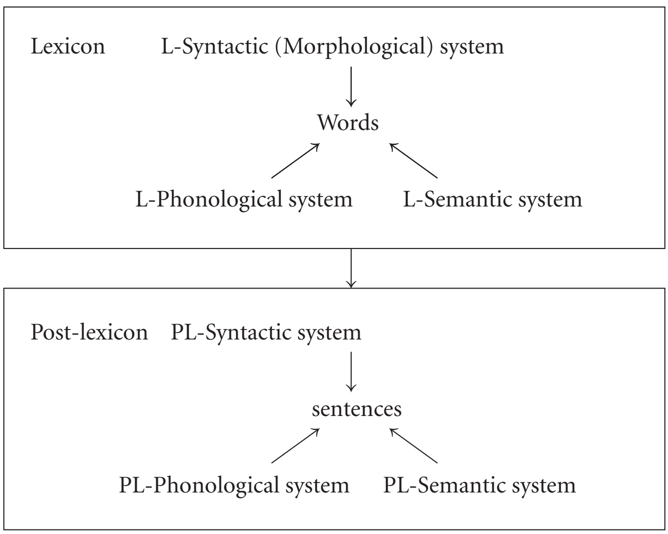
In our view, the heart of the phonological component of a language consists of a set of constraints that evaluate the well-formedness of the phonological representation of linguistic expressions in that language. Constraints can be universal (in which case we could call them principles) or they can be language-specific in the sense that they represent choices from a small set of universal options (in which case we speak of parameters, or parametric constraints). Effectively, this implies that we take phonology to be a constraint-based system, as opposed to the kind of rule-based system that was advocated in Chomsky and Halle (1968). As pointed out in Mohanan (2000), the important distinction here is not the use of terminology ('rule' versus 'constraint'), or the format for expressing phonological propositions (using or not using 'arrows'[3]). The important distinction is between a procedural and non-procedural (declarative) conception of grammar. We take the defining characteristic of a procedural conception to be the use of more than one level of representation.

Non-procedural phonologies (or proposals moving in that direction[4]) have been around in phonology (and linguistics) for a long time (for an historical perspective in phonology see Paradis & Prunet 1993; Scobbie 1997; Bird 1990, 1995). OT instantiates a particular version of this kind of approach with the specific property that, in OT, constraints are violable (or 'soft').

Our own approach is not strictly non-procedural, however, in that we adopt the central claim of Lexical Phonology (which is, in fact, a classical claim, going back to the very beginning of phonology[5]) namely, that phonology is sensitive to a fundamental difference between two types of linguistic expressions: words and sentences (Kiparsky 1982, 1985, 2000). In accordance with that division, there are, then, two sets of well-formedness constraints: a lexical (word-level[6]) set and a 'post-lexical' (sentence-level) set. This, of course, introduces a procedural element in the theory. We do maintain that within each of these two levels, the phonological system is non-procedural and declarative.

In general, we accept the view, defended in Jackendoff (2002), and his earlier work, that the grammar consists of three autonomous (generative) components and sets of interface relations (associations) between the outputs of these components (and grammar external components). We recognize these three components both at the word and the sentence level. An expression (word or sentence) is well-formed if its syntactic, phonological and semantic structures are well-formed and the interface conditions allow the three to be associated. The diagram in (1) shows the organization of the grammar; the interface relations that hold between all subcomponents are not indicated:[7]

(1) The grammar as a checking device




One question that might come up at this point is: where do the words and sentences come from in the first place, such that they can be checked for well-formedness? In generative grammar, the usual answer is that, in each major module, the syntactic system generates the structures. In this view, syntactic 'checking' is implied such that the generative system only produces syntactically well-formed expressions. In this syntactic-centric view, the other two systems (phonology and semantics) have a checking (sometimes called 'interpretative') role only.

In the spirit of Jackendoff (2002), we reject this syntactic-centric view, replacing it by the idea that the expressions that are checked are *all the possible combinations*[8] of L-primes and PL-primes, respectively. The L-primes are **morphemes**, packages of a chunk of phonological structure, a chunk of semantic structure and a syntactic category label.[9] The PL-primes are **words** as allowed by the L-module (the Lexicon).

Each of the six subcomponents in (1) characterizes a finite compositional system, i.e. a system comprising a finite set of primitives and a finite set of constraints on modes of combination, together defining an infinite set of expressions:[10]

(2) ALL GRAMMATICAL SYSTEMS HAVE THE SAME MACRO DESIGN

finite set of **PRIMITIVES** → well-formed **EXPRESSIONS**

finite set of COMBINATORIAL **CONSTRAINTS** → well-formed **EXPRESSIONS**

Adjustment ↑

In (2), we have added one extra step: *adjustment*. Adjustment may seem a necessary step if basic primitives may vary in terms of their 'surface' appearance, depending on the environment that they appear in. Think, for example, of the phenomenon of allophony (with or without allophonic overlap[11]). Traditionally, phonological primitives are subsegmental (subphonemic) units called *features*. The combinatorial constraints ("segment structure conditions") define an inventory of phonemes, abstracting away from what is called allophonic variation. Hence, once phonemes occur in sequential combinations, allowed by the next layer of combinatorial constraints ("syllable structure conditions"), an extra step is required that accounts for the allophonic appearance of phonemes:

(3) **L(exical)**-PHONOLOGY

finite set of primes[12] → well-formed L-phonological expressions

finite set of constraints → well-formed L-phonological expressions

allophony rules ↑

We follow the lead of Government Phonology (GP) in assuming that all allophonic variation is due to the fact that phonological elements may not be licensed in all positions, assuming that non-licensed elements are not phonetically interpreted.[13] In other words, there are no adjustment rules in this model. In the diagram in (3), one might replace 'allophonic rules' by 'phonetic interpretation'. Let us discuss an example. It is well-known that, in English, there is an allophonic alternation between voiceless unaspirated stops and voiceless aspirated stops. Our analysis would go as follows. Firstly, the combinatory constraints dictate that the element 'stop' must co-occur with either the element 'aspiration' or the element 'voice'. This defines /$p^h$/ and /b/ as possible expressions, while it disallows /p/. Then, a second constraint indicates that the aspiration element is not licensed in certain contexts, allowing it in the onset of stressed syllables (unless preceded by /s/) only. Thus /$p^h$I$t^h$/ will be interpreted as [$p^h$It]. We are not saying that the aspiration element is deleted; it is just not interpreted. (A 'traditional' analysis takes the underlying phoneme to be unaspirated /p/, thus requiring an aspiration rule.)

At this point, one might ask how we account for 'allomorphy', i.e. the fact that morphemes can be manifested in different ways depending on the context that they appear in. Our position is that allomorphy is also the result of 'phonetic interpretation', and thus is always due to the fact that certain elements of (segments or) morphemes are not interpreted because they cannot be licensed. After all, allomorphy simply results from allophony if the conditions for one variant happen to be created in a morphologically derived context.[14]

It will be obvious that the approach taken here places severe restrictions on the kinds of allomorphy that the phonology can account for. Indeed, certain types of alternations cannot be accounted for. An example involves 'velar softening' in English (*electric* - *electri*[*s*]*ity*). It has, of course, long been argued for (most forcefully in Natural Generative Phonology; Hooper 1976) that such cases of allomorphy are completely random from a synchronic phonological point of view. We propose that such random allomorphy, in as far as it necessitates a "rule" that substitutes [s] for /k/ in case of -*ity*, is a part of the morphological rule that introduces the relevant affix (cf. Strauss 1982). Another case in point is the rule of 'learned backing' in French (Dell & Selkirk 1978).[15]

In (3), we refer to the set of L-phonological expressions. What type of expressions do we envisage here? Firstly, let us say that expressions occur in a hierarchy of types:

(4) elements (finite set)
|
X-type expressions ('segments')
|
Syllabic expressions (onsets and rhymes)
|
L-phonological words[16]

Thus, elements can be combined forming an expression type that comes close to the traditional notion of 'phoneme'. We call these expressions *segments*, because, indeed, elements are the real 'phonemes' (in the sense of being phonological primes), whereas segments are not necessarily traditional phonemes (cf.

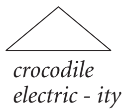
our example of aspiration above).[17] The term 'X-type expression', furthermore, refers to the fact that segments are linearly organized in terms of an X-skeleton. X-slots are organized into the syllabic constituents onset and rhyme, where each constituent is maximally binary in GP. Again, following GP, we do not assume that onsets and rhymes are organized into 'syllables', even though onsets and rhymes occur in packages, in which the onset is a dependent of the rhyme.[18]

Both onsets and rhymes enter into relationships among themselves, called inter-onset relations and inter-nuclear relations, which, as in the case of onset-rhyme relations, are not regarded as establishing constituents.[19]

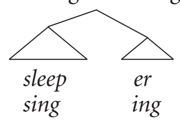
The expression 'L-word' warrants a more detailed discussion. This is a type of unit within which onsets and rhymes form a close-knit organization that is totally independent from 'outside structure' (i.e. L-syntactic ('morphological') structure). Corresponding terminology for this domain would be the notion 'non-analytic domain' (as introduced in Kaye 1995), 'word level' (as used in Borowsky 1993), or the 'α-domain' (as discussed in Inkelas 1989).

In many, perhaps all, languages, morphological products fall into two classes as far as their phonological behavior is concerned. An 'inner class' (formed by affixes that attach to non-derived stems, or stems derived by members of this same class), where the morphological products behave just like underived words, and an 'outer class'. The idea is that words that are derived by affixes of this inner class and words that are underived form L-phonological words (i.e., *L-words*). Then, there is an 'outer class' of affixes, forming L-syntactic ('morphological') expressions that the phonology treats as consisting of an L-phonological word contained within a larger domain. Outer class affixes, in our view, include inflectional affixes, which, perhaps, constitute the most typical members in this class.[20] Finally, compounding appears to deliver combinations of L-phonological words. To capture these three cases, we propose to extend the terminology for L-phonological expressions in the following way:

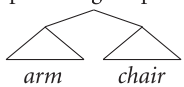
(5) a. L-phonological word




b. L-phonological clitic group




c. L-phonological phrase




The notion 'L-phonological clitic group' corresponds to Kaye's (1995) 'analytic domain' ([[X] Y]) and Inkelas' (1989) 'β–domain'. It is possible that the 'L-phonological phrase' also applies to words that are derived with certain types of 'heavy' prefixes and suffixes, in English and other languages. Evidence in favor of the L-phrase domain would require the identification of phonological generalizations that pertain to this domain exclusively. As far as we can see, the only reason for postulating the L-phrase domain lies in the fact that compounds have a characteristic stress pattern. Stress, being the exponent of phonological constituency, indicates that the members of compounds form a constituent.

These different types of phonological expressions, notably (5b) and (5c), are, effectively, projections from L-syntactic (morphological) structure. On the whole, this proposal is of course very similar to the classical SPE-treatment of the relationship between phonology and morphology, and its later development in Kiparsky's Lexical Phonology (Kiparsky 1982, 1985), based on Siegel's (1984) and Allen's (1978) models of the interleaving of morphology and phonology. Kaye (1995), however, points out that the distinctions in (5) cannot always be predicted from affix classes, or affixation versus compounding. For a specific word-formation process, some complex forms may be analytic, whole others (due to 'lexicalization') may behave as non-analytic. We return to the issue of mismatches between L-syntactic and L-phonological structure below.

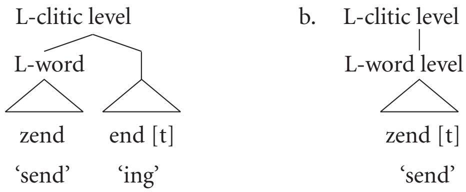
Assuming, then, that the higher levels in the L-phonology, involve at least L-words and L-clitic groups, we need to establish phonological constraints for each such domain. It is clear that the L-words form the domain of what are usually referred to as *phonotactic constraints*, i.e. constraints on combinations of consonants and vowels into onsets, rhymes and feet. In addition, this domain is relevant for phonological constraints that are responsible for 'allophonic (automatic) alternations'. For aspiration in English, we cannot establish whether its domain is the L-word or the L-clitic group simply because L-clitics (being many of the so-called 'Level II' suffixes such as *-er, -ing, -s, -ly* and inflectional suffixes) do not contain unaspirated voiceless stop consonants. However, Borowsky (1993) discusses a host of generalizations from English and other languages that involve automatic generalizations that are bound to the L-word domain.[21] Consider the following words:

| | | | | | | |
|---|---|---|---|---|---|---|
| (6) | a. | cycle | b. | cycling | c. | cycl-ic |
| | | meter | | metering | | metr-ic |
| | | center | | centering | | centr-al |

In *cycle* we find a final 'syllabic liquid', which also occurs (in some dialects) in *cycling*. *Cyclic*, however, has a non-syllabic liquid. The constraint that enforces a final, post-consonantal liquid to be syllabic, can apparently apply in the case of *cycling*, which, Borowsky argues, indicates that *cycling* properly contains the phonological domain (in her terms: the word level) that is relevant for this constraint. In *cyclic*, the /l/ is not domain final, which indicates that this word does not properly contain a word level, but rather constitutes a word-level unit.

A complication in this case with respect to non-derived words ending in *-er* (such as *meter* in (6) above) is that the final /r/ of the agentive suffix *-er* (part of the L-clitic domain) can also be pronounced as syllabic, for example in *eat-er*. This indicates that the L-clitic domain is also subject to the same constraint regarding post-consonantal liquids as found in the L-word domain.[22] If this pattern is typical, it may indicate that the L-clitic group is a recursive version of the L-word, and that constraints that are relevant to the deepest L-word may apply to recursive L-words as well. It remains to be seen whether certain constraints are relevant to the deepest L-word or the recursive L-word only. Borowsky, in any event, concludes that, in English, there are no specific phonological generalizations that belong to the latter domain (what we here call the L-clitic group). However, we believe that this conclusion does not hold in general. Consider 'final devoicing' in Dutch.[23] We take this phenomenon to involve an L-phonological constraint that holds at the L-clitic group level and not at the L-word level. We propose that the relevant constraint states that the voice element is not licensed at the right edge of the L-clitic domain. If, for example, we assume that, in Dutch, the present participle suffix *-end* is an L-clitic, we clearly see that 'final devoicing' is relevant at not only the L-word domain:

(7) a. b.




As (7a) shows, voicing is licensed at the right edge of the L-word, as long as this right edge is not also at the same time the right edge of the L-clitic group.[24]

Thus, we believe that the L-clitic group (or: recursive L-word) *can* form a domain for phonological constraints that are specific to this domain.[25] Another set of examples for constraints that specifically bear on the L-clitic group is the set of *vowel harmony constraints* that apply in all languages where the domain of harmony subsumes inflectional affixes.

A final point with reference to the L-clitic group is the following. In the majority of cases, this domain will be morphologically complex. However, morphological complexity is not a requirement. It has been known for a long time that words may end in consonant coronal clusters that do not seem to 'fit' in the L-word domain in as far as the cluster represents an unusual kind of 'coda' that, apparently cannot occur word-medially, for example as in the word ***first***. Traditional analyses (e.g. Fudge 1969, 1987) represent these coronal clusters as 'appendices'. We suggest that these 'appendices' are L-clitics, even though they do not correspond to synchronic suffixes. Possibly, initial /s/, especially in combination with branching onsets (as in /spl/, /spr/) can be analyzed as an L-preclitic. In other cases, monomorphemic words contain internal 'superheavy syllables', which are severely limited in their distribution, mainly occurring in 'word-final' position (for example, words like *mountain*). Also, monomorphemic words may have the appearance of compounds without being synchronically analyzable as such. Mismatches of this sort are expected given that the distinction between L-words and L-clitic groups is a purely phonological one, even though there are obvious generalizations with respect to the alignment of these domains and L-syntactic (morphological) structure.

## 2.2 Repairs?

Our approach bears a certain resemblance to so-called *constraints-and-repairs* models. The idea of a constraints-and-repairs (CR) approach has been put forward in, among others, Stewart (1983); Singh (1987); Calabrese (1986, 1995, to appear) and Paradis (1988).[26] In comparison to the rule-based approach advanced in Chomsky and Halle (1968, SPE), one might say that the CR-approach breaks up the 'rules' that were used in classical generative phonology, into two parts:

(8) Classical SPE rule

$A \rightarrow B / C – D$

Repair — Constraint

$A \rightarrow B$ — *CAD

Once these two ingredients are teased apart, it becomes possible that one constraint may be 'served' by a variety of repairs that act on different kinds of violations of that constraint. A situation like that cannot be dealt with in a satisfactory manner in a rule-based approach without the loss of generalization, since several rules would necessarily repeat the statement of the relevant constraint, a situation referred to by Kisseberth (1970) as a 'conspiracy'.

Repairs are said to come into action when a combination of morphemes, which is well-formed from the viewpoint of the L-syntactic system and the L-semantic system, violates a constraint of the L-phonological system, either holding at the L-word domain, or at the L-clitic group domain. One might ask whether, analogous to L-phonological repairs, there are also L-morphological or L-semantic repairs? If this is not the case, then it follows that morpheme combinations that violate L-syntactic or L-semantic constraints are apparently **rejected**, rather than repaired. This lack of analogy between L-phonology and the other two L-components is unexpected in a view of grammar that attributes the same role to all subsystems, viz. that of a checking system.

In our approach to phonology, the lack of analogy does not exist. Our claim is that there are no phonological repairs. Phonological representations are interpreted in accordance with the constraints that hold for each domain. It is not relevant whether or not the domain is morphologically complex or not. L-words can be simplex or complex and, as argued above, the same holds for L-clitic groups.

Let us consider an example that involves an alleged conspiracy in some languages:[27]

(9) $*[V\{V,C\}C]_{\text{rhyme}}$ "no superheavy syllables"

A constraint of this type is relevant in Yawelmani, where it seems to motivate two rules (in traditional rule-based analyses): vowel shortening in case the violation is VVC (=> VC), and epenthesis in case the violation is VCC (=> VCVC). In a rule-based analysis, the constraint is encapsulated (and thus repeated) in two rules:[28]

(10) a. VVC => VC (shortening)
b. VCC => VCVC (epenthesis)

In a constraints-and-repairs model, the constraint is stated once (as in 9), while the repairs just say:

(11) a. delete V
b. insert V

The different violations will uniquely select the appropriate repair, assuming that the output of repairs must be well-formed structures (or at least, must not be worse than their input). The approaches of Paradis and Calabrese (cf. references above) offer analyses in this type of framework, addressing many issues of detail.

Another argument for separating constraints ands repairs can be based on *cross-linguistic* conspiracy effects, i.e. cases where different languages choose different repairs for what appears to be the same constraint. Thus, the idea is that for a specific violation of the constraint in (9), say VVC, different repairs are possible in different languages, for example:

(12) a. delete V (VVC => VC)
b. insert C (VVC => VCVC)
c. delete C (VVC => VV)

Steriade (2000) has recently pointed out that it is, in fact, not at all common to find that different languages select different repair solutions for what seems to be the same constraint violation. One of her examples involves the following constraint:[29]

(13) *[obstruent, voice] #

Of all the conceivable repairs (insert a vowel word-finally, delete the entire consonant, change the consonant into a sonorant and so on), the one that is always found is:

(14) Delete voice

Our own approach concurs with Steriade's findings. Indeed, since our model has no repair rules to begin with, languages could not possibly differ in having different repair rules. In our model, "repair" is embodied in the way in which phonological structure is interpreted. Following the original proposals in Government Phonology, we pursue the ideal that the interpretation of phonological constellations is entirely universal. Differences between languages cannot result from different 'repairs' of the same constraint. Rather, the well-formedness of phonological representations is dependent on universal constraints ('principles'), as well as parametric constraints. The latter lie behind differences among languages.

How does OT handle 'repairs' if it only has constraints? It is assumed by OT-proponents that the situation in (12) is a realistic one: different languages *can* respond differently to one and the same constraint. OT handles repairs by incorporating 'faithfulness constraints' that penalize discrepancies between the input and the output, for example:[30]

(15) a. *delete C (read as 'do not delete C')
b. *insert C (read as 'do not insert C')

By confronting the constraint set with all logical alternative 'repairs' (the candidate set) for a given input, and by postulating an extrinsic ordering on the faithfulness constraints, OT achieves the result of selecting a unique output for any given input. The lowest ranked 'anti-repair constraint' specifies the repair that a language will allow, assuming that there is a higher ranked well-formedness constraint that enforces the violation of this repair constraint:

(16) a. *[VVC]$_{\text{rhyme}}$ » *insert C » *delete (VVC and VCVC lose, VC wins)
b. *[VVC]$_{\text{rhyme}}$ » *delete V » *insert C (VVC and VC lose, VCVC wins)

Van der Hulst and Ritter (2000b) point out that the adoption of 'anti-repair constraints' forms one of the two important reasons for introducing parochial ordering into the set of constraints.[31] It should be clear at this point that parochial ranking is what separates our approach from OT, nothing else.[32]

## 2.3 On the lexical/post-lexical distinction

With reference to the post-lexical component, (2) is instantiated in the following way:

(17) PL-PHONOLOGY

finite set of L-phonological expressions[33] → well-formed PL-phonological expressions
finite set of constraints → well-formed PL-phonological expressions
phonetic interpretation → well-formed PL-phonological expressions

When words are combined, all sorts of ‘processes’ seem to apply that cause variation in their phonological shape, dependent on neighboring words or position in the phonological hierarchical structure. Analogous to what we proposed for the L-phonology, we postulate that in this case too variation is due to phonetic interpretation.[34]

Now, whereas the L-phonological system characterizes exactly one L-representation (for a given idealized speaker), every such speaker is capable of pronouncing utterances in many different ways, depending on factors that relate to tempo and style of speech, and other sociolinguistic factors, which must be built into the constraint system.[35] A PL-model that incorporates some kind of gradient parameterization of constraints in order to account for stylistic variation has not been designed, as far as we know. In addition to needing gradient parameterization, it seems likely that the ‘combinatorial syntax’ of the PL-phonology is different from that of the L-phonology. It may be the same to a large extent, but we believe that certain differences must be assumed. For example, the notion of ‘ambisyllabic consonant’, and perhaps in general the idea of overlapping constituents and improper bracketing seems to belong exclusively to the post-lexical level.[36] In short, the PL-system is largely a terra incognita.

Let us now briefly address the question as to why we wish to adopt the lexical/post-lexical distinction. After all, this move introduces a procedural element in the theory. There are several reasons, but we focus here on Kiparsky’s proposal that the lexical/post-lexical distinction is necessary to account for certain cases of ‘legitimate’ opacity. The separation of a word and sentence constraint set implies the possibility of the latter obscuring the former, creating opacity effects. Thus a lexical constraint barring initial clusters may be overridden by a post-lexical constraint that allows such clusters. We follow Kiparsky (2000) in adopting the hypothesis that many ‘legitimate’ opacity effects arise in this way. Why is this crucial? It seems that constraint-based approaches (including OT) are fundamentally unable to deal with opacity effects, without allowing some kind of procedural ordering into the system. Kiparsky’s approach limits extrinsic ordering to ordering between entire constraint sets. We believe that this is the most principled way of accounting for certain (but not all[37]) opacity effects. In addition, differences between lexical structure and post-lexical constituent structure may produce what van der Hulst (to appear a) has called *Structure Paradoxes*. Note that such paradoxes are simply one instance of the broader phenomenon of opacity.

We might note that the ordering of constraints implies some kind of opacity as well. After all, if a constraint A outranks another constraint B (and A and B describe incompatible well-formedness properties), A *obscures* the presence of B. For example, if a constraint that describes that stress is initial (B) is outranked by a constraint describing that heavy syllables must be stressed (A), any word that starts with a light - heavy sequence will cause the initial stress constraint (B) to be opaque. The fact that constraints (if outranked) can be violated in the output makes them soft constraints. Recall that declarative constraint-based systems (such as Scobbie 1997; Bird 1990, 1995) disallow this kind of opacity. To give it a name, let us refer to the opacity of constraints, due to ranking, as *non-procedural opacity*.[38] The kind of opacity that Kiparsky tries to avoid (and that motivated extrinsic ordering in SPE) is of a different kind that we will call *procedural opacity*.[39] Proponents of OT find it hard to accept, but constraint ordering *cannot* produce procedural opacity effects.[40]

So why, one might wonder, does Kiparsky (2000) adopt the extrinsic ordering of constraints, if it doesn't help to account for procedural opacity? The answer, we believe, has been partly given above. By introducing competing 'repair' constraints into the constraint set, ordering has to be imposed to select the correct output. There is, moreover, a second reason for why OT adopts ordering, resulting from the rejection of parameters. We will discuss this in Section 4.

We conclude this section, by recapitulating our position on the use of procedural mechanisms.[41] We make a three-way distinction:

a. We have adopted a two-level (lexical/post-lexical) approach to phonology. Since both levels are ordered (as a consequence of the organization of the grammar), our approach to phonology is procedural in the sense that the constraints on sentences take precedence over (i.e. come after) the constraints on words. This kind of ordering is totally universal in that it follows from the organization of the grammar as a whole.
b. Secondly, our model does not make use of extrinsically ordered rules that create intermediate steps (within the lexical or post-lexical level).
c. Thirdly, our model does not appeal to parochial ordering of constraints as used in OT. This ordering does not create intermediate levels.[42] However, specific constraint pairs may be ordered in terms of an 'elsewhere relation' in our model, as discussed in our analysis in Section 3.3 below (also cf. Ritter 1995 for such a discussion with respect to resolving parameter conflicts).

Thus, the approach that we explore is procedural in the sense of (a), but it rejects (b) and (c). Proponents of Government Phonology usually do not like to make the lexical/post-lexical distinction (Kaye 1995). This is largely a 'programmatic' claim since little work on sentence level expressions has been done in this model.[43] Dependency Phonology, however, incorporates the distinction explicitly. Standard OT (Prince & Smolensky 1993) rejects all three procedural elements, although it has something else that looks a lot like (b): the evaluation procedure is based on an input/output distinction. OT, in Kiparsky's approach, adopts (a). Declarative Phonology (Scobbie 1997; Bird 1990, 1995) rigidly rejects all three procedural elements, although one can conceive of a declarative phonology that adopts the lexical/post-lexical distinction (Coleman 1995a).

## 3. Head-driven Phonology

The preceding section presented the general outlines of a two-level (lexical/post-lexical), (declarative) constraint-based phonological model. In this section, we will discuss a specific theory that is compatible with these outlines.

### 3.1 The parallelism between phonology and syntax

A model that, in our mind, comes close to incorporating all the hypotheses that we have adopted above is Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990). This model is constraint-based (incorporating principles and parameters). There is no extrinsic ordering of any kind.

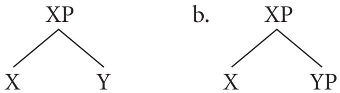
In most other constraint-based approaches, the rejection of intermediate levels has lead to extremely 'surface-oriented' and 'construction-specific' constraints that shy away from anything that is not 'phonetically present' in the speech signal (cf. Natural Generative Phonology, Declarative Phonology). Even in OT, where constraints can be obscured in the output, there is a tendency to stick to the "phonetic facts". Government phonology is different from these surface-oriented approaches. Its presupposition is that phonology characterizes a cognitive system (cf. Kaye 1989). As such there is not even the possibility of referring to 'phonetic events'. On the contrary, the kinds of constraints that one finds in Government Phonology are general and refer to rather abstract properties of phonological representations. In doing so, Government Phonology has taken the explicit view that phonology is *not different* from syntax in allowing its expressions to contain abstract units that are essential in accounting for the well-formedness of expressions without necessarily receiving a phonetic interpretation.[44]

A second extremely attractive feature of this model involves its reliance on *head-dependency relations* at all levels of the phonological representation.[45] In adopting this feature, Government Phonology has been preceded (and influenced) by the model of Dependency Phonology (Anderson & Ewen 1987).

The parallelism between syntax and phonology that we signaled above, goes much further than simply allowing 'abstract units'. The fundamental hypothesis of the government/dependency program is what Anderson and Ewen (1987: 283ff.) refer to as the *Structural Analogy Hypothesis*. According to this hypothesis, we expect, all things being equal, that the cognitive systems of syntax and phonology are organized in identical ways.[46] This view is in contrast with that of Bromberger and Halle (1989) (and many others) according to which 'phonology is different'. For Bromberger and Halle this means that phonology can have extrinsically ordered rules, while syntax does without such a mechanism. We do not subscribe to this extreme version of modularity. Maintaining that phonology and syntax are different modules, we pursue a highly parallel organization, subject to general principles of UG, for both modules.[47]

There is also a terminological consequence to the parallelism between phonology and syntax. Firstly, the term 'syntax' just means 'putting things together'. The modules of phonology and syntax both rely on a finite set of primitives and a finite set of rules in order to characterize their expressions. Thus, both modules are 'syntactic'. We claim that the 'syntax' of phonology and the 'syntax' of syntax share interesting properties such as binarity of structure and headedness.[48] Do the parallels extend further? They do, perhaps, in the more general claims of UG, for example, in that the projection of underlying structures cannot be tampered with in an ad-hoc fashion and that units (be they syntactic or phonological) must be licensed in their positions (especially those positions that are empty). It might also be asked whether there are phonological parallels to transformations. In the old days of generative grammar, the answer was simply affirmative: phonological rules were the analogues to syntactic transformations. The phonological model that we support does not have transformational rules. But then again, syntax can be done without transformations as well. In any event, the hypothesized parallelism does not imply that phonologists must blindly 'copy' the theories of their syntactic colleagues, or vice versa. Both phonologists and syntacticians pursue the best possible theory for their domain. They can look at each other's results in an heuristic spirit and import concepts back and forth, but the structural analogy hypothesis is ultimately an empirical claim. Moreover, it does not predict that both components are fully identical. A common set of principles and parameters may result in different structures in response to the fact that both domains start out with different sets of primitives (cf. van der Hulst 2000 for extensive discussion). Lastly, we should not be reluctant to see significant differences between the 'syntax' of phonology and the ‘syntax’ of syntax. One such difference might involve the notion of recursion. It seems to be a typical property of phonological constituency that different constituent types are arranged in ‘strict layers’, where each unit dominates one or two units at the lower layer, as in (18a). Syntactic structures are quite different and this is due to the fact that constituents can contain a constituent of the same type as a dependent of the head, as in (18b):

(18) a. b.



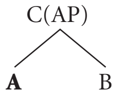

Dresher and van der Hulst (1998) refer to the relationship in (18a) as αα–dependency, and that in (18b) as αβ–dependency, simply to indicate that in the former case, head and dependent are of the same type (i.e. X-slot, syllable, foot and so), whereas in the latter case head and dependent are of different types. The lack of recursion in phonology seems to imply that the set of phonological expressions is finite.

## **3.2** The expression of asymmetry in phonology

The Head-dependency Principle (HD) is the fundamental expression of *asymmetry* in phonology. The central claim is that phonological representations are driven by an asymmetrical relationship between the units. In fact, we believe that the idea of headedness is so fundamental that we have chosen to refer to our own views, which strongly build on Government and Dependency phonology, as *Head-Driven Phonology* (HDP).[49] The fundamental principle can be stated as follows:

(19) The Head-dependency Principle (HD)
All phonological relations involve one head and at most one dependent

In our earlier work (van der Hulst & Ritter 1999), we also formulated a Binarity Principle, which stated that phonological units are maximally binary. One might argue that binarity is implied by the above HD-principle. Note that (19) does not say that all phonological units are necessarily binary. An onset, for example, may contain one consonant.[50] A phonological representation is a network of binary relationships, and it will be necessary to distinguish a few different types of relationships. Some of these relationships correspond to structural notions such as sisterhood and dominance:

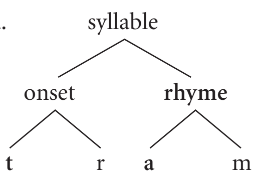
(20)


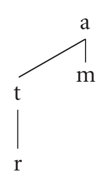

Assuming that the bold-face A is the head, a constituent implies a HD-relationship between C and A (A is the head of C; C is an A-phrase or AP). At the same time there is a HD-relationship between A and B (A is the head of B; A governs, or licenses B). One might, in fact, conflate these two head roles of A by replacing (20) by (21)

(21)



This assumes that the linear relationship between A and B (if present) is accounted for separately, in terms of 'linearization rules'. In Dependency Phonology, representations such as (21) have always been used for subsegmental structures since subsegmental relationships between phonological primes (features or elements) clearly do not involve linear order. But, even in domains where linear order is apparently relevant, such as syllable structure, it can be maintained that the phonological structures do not express linear order (which is left to separate linearization rules). Consider an English syllable like *tram*. Assuming that this syllable consists of an onset (tr) and a rhyme (am), we could adopt a traditional representation as in (22a), or a different representation as in (22b):

(22) a.



b.



Given that (22b) is one constituent, it is predictable, because onsets universally precede Rhymes, that the (a,m) unit follows the (t, r) unit in order to constitute a well-formed syllable. Thus, (22b) should not be taken to express linear order. We should also note that in the case of syllabic constituents (onset, rhyme), the location (in terms of linear order) of the head is universally fixed: onsets and rhymes are left-headed.[51] Thus, the structure in (22a) can be taken as a linearization of the structure in (22b). At the foot and word layer, however, it appears that the orientation of the head is not fixed (i.e., it is parameterized).

The relevance of the head – dependent relation as part of constituent structure goes beyond determining the *syntagmatic* relation of linear precedence. In addition, heads and dependents differ in what they can dominate; we can call this the *paradigmatic* side of constituent structure. Here we mention an important principle that constrains the nature of structural paradigmatic relationships (cf. Harris 1990; Dresher & van der Hulst 1998):

(23) The Asymmetry-Complexity Principle[52]
Dependents cannot be more complex than their heads

This principle captures a fundamental property of representations. It states that in cases where head and dependent positions allow different units, head positions allow a greater array of contrasts than dependents. This has obvious distributional consequences. For example, head syllables allow a greater array of vowels than dependent syllables, or head syllables allow branching onsets, whereas dependent syllables do not. Another example is that heads of onsets allow a greater array of consonants than dependents of onsets do.

Government phonology has also made a case for head-dependent relations that do not refer to units that form a constituent in the traditional sense. We already saw this in the claim that the relationship between onsets and rhymes is not one of constituency. Another case in point is the Coda Licensing Principle (Kaye 1990):

(24) Coda Licensing
A coda must be followed by an onset that can license it

According to this principle, no word can end in a coda, and even word-medial codas are restricted in that the coda - onset sequence (called the interlude) must be such that the onset consonant can license the coda. Roughly, in order to have this potential, the onset consonant cannot be more sonorant than the coda. This principle accounts for the fact that whereas the sequence *n-t* is generally a well-formed 'interlude' (in languages that allow codas in the first place), *t-n* is not. Coda licensing is a head - dependent relationship in which the onset is the head, despite the fact that interludes do not form constituents in the traditional sense.

So far, head - dependency relations have been made 'responsible' for licensing units in the phonological structure. In particular, the claim is that dependent units must be licensed to occur by a head. In addition, we may need constraints that refer to structural positions such as 'edge of' a domain. It remains to be seen whether such *edge constraints* can be reduced to constituent or interconstituent head-dependency relations.[53]

We conclude that phonological representations are constellations of head-dependency relations that are subject to universal constraints (principles) and more specific constraints (parameters).

### 3.3 The role of interpretation in accounting for alternations

In our approach, 'phonetic interpretation' plays an important role. The result of interpretation at the L-level is, in our view, a *phonological* representation that forms the starting point for the post-lexical phonology. We assume that L-units that are not licensed, and thus not phonetically interpreted, are invisible to the PL-system, where combinations of L-words are subject to a (gradient) constraint system that characterizes the PL-level. PL-representations are also subject to 'phonetic interpretation'. The result of that can be phonetically implemented.

In this section, our goal is to show the important role of interpretation (at the L-level) in apparent cases of 'insertion' (producing vowel – zero alternations) in the lexical phonology. We also refer to van der Hulst and Ritter (2000a) for a discussion of the role of phonetic interpretation in accounting for alternations that involve apparent 'deletion' and resulting opacity effects.

Here we present a sample analysis of some phonological patterns in Yawelmani, involving the kind of data that motivated extrinsic rule ordering in SPE. We will show how the data can be accounted for in Head-driven Phonology. The analysis relies heavily on the original proposals made in the context of Government Phonology, but differs in various important respects from the analysis put forward in Kaye, Lowenstamm and Vergnaud (1990).[54] The focus will be on aspects of the data that have led prior researchers to postulate a rule of vowel epenthesis and a rule of vowel shortening, to be applied in that order.

#### 3.3.1 *The relevant data*

Vowel shortening is motivated by the following paradigms (cited from Kenstowicz and Kisseberth (1979: 83 ff.):

| (25) | Nonfuture | Imperative | Dubitative | Future | gloss |
|---|---|---|---|---|---|
| | dos - hin | dos - k'o | do:s- ol | do:s - en | report |
| | lan - hin | lan - k'a | la:n - al | la:n - en | hear |

Roots that invariably have short vowels show that the above roots have underlying long vowels that apparently shorten before two consonants (which amounts to closed syllable shortening in some models).

Epenthesis is illustrated by the following data:

| | | | | |
|---|---|---|---|---|
| (26) | logw- ol | 'might pulverize' | logiw - hin | 'pulverizes' |
| | lihm - al | 'might run' | lihim - hin | 'runs' |

An epenthesis analysis is more likely than a deletion analysis since there are otherwise no CVCVC roots in the language. This is explained if one assumes that there are, in fact, no such roots in the language and that the surface pattern CVCiC is derived through epenthesis. The rule of epenthesis breaks up a cluster of three consonants.

The two rules are extrinsically ordered as shown by the following data:

| | | |
|---|---|---|
| (27) | sonl - ol | 'might put on the back' |
| | so:nil - mi | 'having packed on the back' |

The underlying form for these roots must be CV:CC. A standard derivation of both forms shows that epenthesis must precede shortening:

| | | | | | | |
|---|---|---|---|---|---|---|
| (28) | (a) | so:nl - ol | so:nl - mi | (b) | so:nl - ol | so:nl - mi |
| | Ep. | — | so:nil - mi | Short. | sonl-ol | sonl - mi |
| | Short. | sonl - ol | — | Ep. | — | *sonil - mi |

By standard reasoning both rules must be extrinsically ordered.

### 3.3.2 *An HDP analysis*

In our model, there are no deletion or insertion operations. Following the ground-breaking idea of Government Phonology, vowel – zero alternations are handled by assuming that the input forms contain empty vowel positions that can be interpreted as silent or audible depending on the environment.[55] The crucial generalization is stated in (29):[56]

(29) An empty nucleus can be silent only

a. if it is governed by a following non-empty nucleus (*Proper Government*)
b. if it is final in its L-phonological domain (*Final Licensing*)

The implication of (29) is that a nucleus that is phonologically empty may be silent (if it meets 29a or 29b) or audible (if it does not meet 29a or 29b).[57] In practice, (29) makes it impossible for two adjacent empty nuclei to both be silent.

Let us first consider a form in which empty nuclei are silent:

(30) O N O N O *N* - *O* N O N /logw - ol/ → logwol

l o g w o l

⟵ √

The second nucleus from the left is the locus of the vowel - zero alternation, while the final empty nucleus is 'forced upon us' by the Coda Licensing Principle which disallows final codas. We have argued earlier that in GP, each onset co-occurs with a nucleus, and vice versa; in these packages, onsets are dependents while nuclei are heads. There are no sequences of N's or O's. This, in fact, necessitates postulating both empty nuclei and empty onsets (even when no alternations are present). In addition, a general convention seems needed that allows us to 'ignore' a sequence of an empty nucleus followed by an empty onset (i.e. the italicized sequence NO) (as proposed in Gussmann & Kaye 1993; and Yoshida 1993). We adopt an alternative, suggested by John Harris, which assumes that the morphological concatenation 'overlays' the empty suffix onset over the stem-final onset and the suffix vowel over the empty stem-final vowel such that the resulting representations will not contain the vacuous NO sequence. Thus, (31) replaces (30), showing suppression of the empty NO-sequence:

(31) O N O N O - N O N /logw - ol/ → logwol

l o g w o l

⟵ √

The representation in (31) is well-formed. The left-pointing arrow represents Proper Government (PG), the HD-relation that licenses empty nuclei to remain silent. Hence, in this form, no empty nucleus is interpreted as audible; in other words: there is no 'insertion'. The '√' indicates Final Licensing (FL).

The question might be raised as to whether PG constitutes a constituent. The relationship between an empty nucleus and a following full nucleus is clearly 'foot-like' (as suggested in van der Hulst & Rowicka 1997). Here, we will leave this issue open. Clearly, the foot-like relationships are not responsible for the familiar rhythmic patterns in stress systems which, in our view, result from post-lexical foot structure (cf. van der Hulst, to appear a, b). In this sample analysis we do not discuss word stress in Yawelmani.

Now consider the next form:

(32) O N O N O N O N O N /logw - hin/ → logiwhin

l o g [i] w h i n ⟵ √

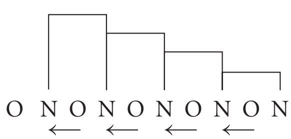
As before, the final empty nucleus is licensed by FL.[58] The penultimate vowel properly governs the empty nucleus to its left. Hence, the empty nucleus *preceding* this properly governed empty one cannot remain silent; consequently, it is phonetically interpreted as the vowel sound [i].[59] The form in (32) tells us that PG 'operates' in a leftward fashion. We must assume this because otherwise the representation in (33) would be a possible candidate:

(33) O N O N O N O N O N /logw - hin/ → *logwihin

l o g w [i] h i n ⟵ √

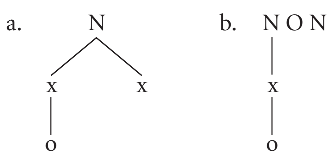
More needs to be said about this right-to-left directional property of PG. We suggest that the directionality of the binary government relationship and the unbounded right-to-left propagation of this relationship form part-and-parcel of the right-headed, right-branching constituency that organizes all nuclei into the L-word domain:

(34)

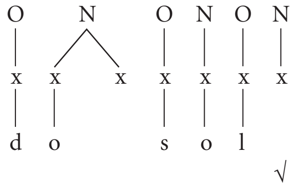



Next we turn to vowel shortening. Long vowels, in general, can result from two different structures, involving either a branching nucleus, or a sequence of two non-branching nuclei (with an intervening empty onset):[60]

(35)



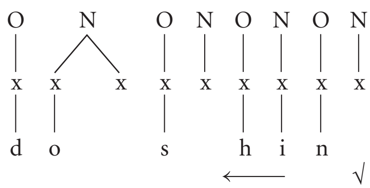

In both cases, the long vowels result from the fact that the second empty position is interpreted with the content of the first position. In the case of Yawelmani, we will show that (35a) allows us to provide an explanatory account of 'vowel shortening', i.e. the failure of the second position to be interpreted as involving the element that is associated to the first position.

(36)



/do:s/ - ol/ → do:sol

Example (36) is a representation of a stem with a long vowel where the second position of the branching nucleus is interpreted with the element associated with the head or first position of that nucleus.

Now consider the form that shows 'vowel shortening'. Here we have a sequence of two empty nuclei and we might therefore expect a lexical foot between them that would force audibility on the right-hand side of the long vowel, an undesirable result:

(37)



/do:s - hin/ → doshin

We now have to explain why the second position of the 'long' vowel cannot be phonetically interpreted in (37), though it is in (36). At first sight, one might argue that the interpretation of the second position of the branching nucleus in (37) is expected because that nucleus cannot be properly governed by the following nucleus that is itself properly governed to remain empty. We must assume that something overrides PG in this case. The GP literature contains a few different proposals. Here we propose our own, leaving the comparison and discussion for another occasion. Recall the principle of Coda Licensing, which demanded that a coda position be licensed by a following onset. We propose that the 'second half' of a long vowel must likewise be licensed. The candidate licensor for the second half of the long vowel is /s/, the immediately following consonant, both in (36) and in (37). The difference between both cases is that in (36) /s/ forms the onset of a non-empty nucleus, whereas this is not so in (37). We propose that onsets of empty nuclei cannot 'coda' license because, so to speak, the empty nucleus that follows the onset (which itself governs the onset) is incapable of giving the onset 'licensing power'. In other words, we propose that the constraint 'Coda Licensing' embodies the constraint that a licensor must be governed by a non-empty nucleus. We will refer to this new formulation of Coda Licensing as 'Revised Coda Licensing'. The idea of governors needing 'licensing power' is not new and there are several GP analyses that appeal to this notion (e.g. Charette 1990). Here its use is extended to a new situation. Thus in (36) the onset that licenses the second position of the branching nucleus is governed by a non-empty nucleus, resulting in the interpretation of that branching structure as a 'long vowel'. In (37), however, the nucleus governing the onset-licensor is empty, resulting in the onset's inability to license the preceding 'coda-like' nucleic position thus yielding the effect of 'vowel shortening'.

The analysis proposed here reveals that a nucleus dependent position is subject to a constraint (Revised Coda Licensing) which takes precedence over Proper Government. We assume that this precedence relationship follows from the Elsewhere Condition because whereas PG makes reference to empty nuclei in general, Revised Coda Licensing refers to *dependent* empty nuclei. Thus, we see here another instance of the relevance of head-dependent asymmetries.

## 4. Problems with parameters?

In this section, we will propose (or rather make explicit) how the setting of parameters is dependent on head-dependency relations. Our proposal will be that the setting of a parameter can differ for heads and dependents. Thus, a language *can* show evidence for opposite settings of a parameter. In the OT literature, it has been argued that the all-or-nothing approach of parameter theory fails to explain 'emergence of the unmarked', which is evidence for the unmarked setting, in a system that also shows evidence for the marked setting of a parameter. Our proposal embodies an answer to this kind of criticism.

Let us start this section, by recapitulating why we refer to our approach as 'constraint-based' if it is using principles and parameters. Firstly, we take principles to be constraints that hold true for all languages. Then, there are statements that specify choices that languages make from a small set options; these are the (set) parameters. Effectively, then, both principles and parameters (once provided with a value) *are* constraints in that these statements specify the well-formedness of phonological representations without any procedural elements.

Parameters, then, can be thought of as binary constraints. Fixing a parameter's value, turns a parameter into a (unary) constraint. This does not mean that languages differ in having 'positive' constraints or 'negative' constraints, corresponding to the positive or negative settings of the parameters. Both settings produce the same kind of constraint (which often can be formulated in a positive or negative way). Take, for example, foot formation, assuming that feet can be left-headed or right-headed, we see this as a parameter. A language that has LH feet can be said to have LH feet (a positive statement) or we can say that it disallows RH feet (a negative statement). Either way, the language has a simple unary constraint.

Optimality theory has no parameters. This means that it only allows unary constraints from the start.[61] Since all constraints are universal,[62] and since languages may have LH or RH feet, differences between languages necessitate extrinsic ordering of the constraints, which, then, is the second reason for why OT needs extrinsic ordering, the first one being the need to rank the anti-repair faithfulness constraints (discussed in Section 2.2). The two alternative approaches are summarized in (38). PPT stands for 'principles and parameters' theory:

(38)

| | Language A | Language B |
|---|---|---|
| PPT | FootHead (left) | FootHead (right) |
| OT | FootHead (left) » FootHead (right) | FootHead (right) » FootHead (left) |

Cases where the two values of a parameter correspond to two truly independent constraints (i.e. constraints that are not each other's opposite) are difficult to find and we should not be misled by the fact that the two opposite constraints sometimes have unrelated names. The difference between languages that have quantity-sensitive (QS) and quantity-insensitive (QI) feet would be made as follows in the two theories:

(39)

| | Language A | Language B |
|---|---|---|
| HDP | QS (yes) | QS (no) |
| OT | Weight-to-Stress » Uniformity | Uniformity » Weight-to-Stress |

The constraint Uniformity says that syllables must be grouped in a uniform way, forming regular groupings of two syllables. Weight-to-Stress requires that every heavy syllable is stressed, i.e. the head of a foot. The constraint *Uniformity* is violated when syllables are not grouped in a regular alternating fashion, which, of course, happens when Weight-to-Stress interrupts a regular grouping, demanding for example that two adjacent heavy syllables each forms its own foot. Now, it may seem as if the OT account has the advantage of appealing to independent forces, but it is really not obvious why 'uniformity' couldn't be called 'weight-not-to-stress', or why 'weight-to-stress' could go under the heading of 'non-uniformity'.

It seems hard to believe that OT has gained its popularity because of the difference outlined here. In fact, given the exponential power of constraint ranking, one would always adopt the PPT model as the more restricted one. The basic calculation is that adding a parameter to your model multiplies the set of possible grammars by 2. Adding a constraint multiplies the number of grammars by whatever the number of constraints was (plus 1).

Proponents of OT have argued that OT is *conceptually simpler* than a principles-and-parameters approach, especially if this latter approach is combined with repair mechanisms. OT uses only one type of mechanism: constraints. This is possible because parametric choices have been 'translated' into pairs of opposing constraints and the repair rules have also been 'disguised' as anti-repair constraints. We have argued that the 'conceptual argument' comes with a price: parochial ordering. It seems obvious that the conceptual argument of simplicity loses much of its force, if the gain in simplicity is counterbalanced by adding the complexity of ranking.[63]

In this section, we will focus on another, more interesting line of argumentation that OT-proponents have followed in dismissing a PPT-approach. The general point is, according to OT, that a principles-and-parameters theory (PPT) is forced into the position that any parameter, once it is set, must hold for the whole language. This position, the argument goes, cannot be correct because there are situations in which a language shows signs of both values of the parameter. It seems to us that we are dealing here with at least two types of cases, which we will discuss in the next two subsections.

## 4.1 Emergence of the unmarked

In OT, the grammar consists of constraints only, i.e. statements that specify "unmarked" properties of language (well-formedness or 'markedness constraints') and statements that bear on the relationship between input and output (faithfulness constraints). In order to express cross-linguistic diversity, it is claimed that the constraints are ranked differently in different languages. In each language, the specific ranking determines which candidate for some input is optimal. This works as follows. Each candidate will violate one or more constraints, i.e. no candidate is perfect. Since the constraints are ranked, each candidate has 'a highest violation'. Now we only need to compare the highest violations of all candidates and pick out the candidate whose highest violation is lower than the highest violation of all other candidates. That candidate is said to be the optimal candidate.

Constraints whose 'demands' are overruled by a higher-ranked constraint, need not be totally invisible in a grammar. Their relevance may be seen in cases for which the higher-ranked constraint is not relevant. This is called *emergence of the unmarked*.

With these points in mind, McCarthy (1995) claims that, in a PPT-approach, the setting of a coda-parameter to the marked option YES (i.e. allowing closed syllables) and the onset-parameter to the marked option NO (i.e. allowing onsetless syllables) fails to explain why a VCV string is always parsed V.CV. Given that a language allows codas and onsetless syllables, why isn't the parsing VC.V also possible? To address this 'problem' it might seem that the parameter allowing a coda must be made *context-sensitive* in the sense that it is 'no' in the context 'V-V' and 'yes' elsewhere. (Of course, if a language has no coda whatsoever, it is simply 'no' in all contexts.) This looks like a cumbersome and uncontrolled enrichment of the theory of parameter setting. OT does not have a comparable problem because even if a language allows onsetless syllables, the constraint Onset (penalizing onsetless syllables) is still part of the grammar and no matter what the ranking of NoCoda and Onset, the CV.CV candidate will always be optimal:

(40)

| CVCV | NoCoda | Onset | \| | Onset | NoCoda |
|---|---|---|---|---|---|
| CV.CV | | | \| | | |
| CVC.V | * | * | \| | * | * |

However, we submit that this case is not problematic in a PPT approach at all. There is no need to make the coda-parameter context-sensitive in some ad hoc fashion. To explain the universal emergence of the V.CV parsing we only need an inviolable and dominating constraint (a principle), that rules out the VC.V parsing universally. That constraint, as it turns out, is already in the grammar for independent reasons: *Coda Licensing*, proposed in Kaye (1990) and discussed above. Coda licensing disallows VC.V parsings. At the same time it also ensures that onsets are maximized (cf. the so-called Maximal Onset Principle) because, as pointed out in Kaye, Lowenstamm and Vergnaud (1990) and Kaye (1990), an onset can only license a preceding coda if the onset consonant is 'less sonorous' than the preceding coda (cf. Harris 1990 for a formalization of the notion 'sonorant' in terms of complexity). A string VbrV can thus never be parsed as Vb.rV because an /r/ can never coda-license a /b/. Such a string will be parsed as V.brV in a language that permits branching onsets.

We now turn to a second, similar objection to parameter theory that has been made in OT-work (cf. Rosenthall & van der Hulst 1999).

## 4.2 Head-sensitive parameter setting

The type of case that we investigate here has been discussed extensively in Rosenthall and van der Hulst (1999). These authors looked at a variety of cases where 'weight-by-position' (WBP) seems to be context-sensitive. WBP states that codas contribute to weight for the purpose of stress assignment. In a PPT approach, we can see WBP as a parameter (as proposed in Hayes 1989). If coda consonants contribute to weight, the value is positive; otherwise, it is negative. In the simple case, all coda consonants behave the same in a given language. Rosenthall and van der Hulst call attention, however, to languages in which the weight of codas seems to depend on the position of the syllable in the word. Hence they refer to this as 'weight-by-position-by-position'.

A typical language in point is Chugach. In this language, closed syllables count as heavy in word-initial position, whereas they count as light elsewhere. (Syllables with long vowels are heavy in all positions in the word.)

The authors claim that a PPT model has difficulty with expressing this state of affairs without 'fragmenting' the weight-by-position parameter. Essentially, they argue, two separate parameters have to be formulated, one that is context-sensitive (41a iii) and one that states the 'elsewhere' case, as in (41a iv). In OT, on the other hand, the situation can be neatly accounted for in terms of a constraint ranking such that a constraint 'Stress-Edge-left', requiring a stress on the left edge of the word, outranks the constraint that prohibits codas from being weightful (*WBP):

| | | | | |
|---|---|---|---|---|
| (41) | a. | HDP: | i. | Foot (left-headed) |
| | | | ii. | Word (left-headed) |
| | | | iii. | WBP-initial (yes) |
| | | | iv. | WBP-elsewhere (no) |
| | b. | OT: | i. | Foot: left-headed » Foot: right-headed |
| | | | ii. | Word: left-headed » Word: right-headed |
| | | | iii. | Stress-Edge-left» *WBP |

Thus, in contexts where both constraints are relevant, Stress-Edge-left will obscure *WBP, but in word-medial context, where Stress-Edge-left has no relevance, *WBP emerges.[64] The effect is that an initial CVC will count as heavy, while a non-initial CVC will count as light.

The OT-account seems to capture the conflict between a coda 'wishing to be weightless' and stress wishing to be on the edge (requiring the left-edge syllable to be heavy). The PPT account, on the other hand, seems clumsy. In particular, it does not seem to make a connection between the 'initial' part of (41aiii) and the fact that (41ai) and (41aii) together specify initial stress. Can we improve the PPT account in a principled way such that it no longer misses a generalization? It seems to us that we can.

In van der Hulst (1984), the idea was put forward that primary and secondary stress should not be collapsed into one algorithm for accent assignment. The reasons for that view were independent from the kinds of facts that we find in Chugach, but these facts can be taken to support the separation of primary and secondary stress, simply because in this case, the weight sensitivity of the primary stress foot (the left-hand foot) and secondary stress feet appear to be different.[65]

Now, given that the stressed initial syllable, being the syllable with primary stress, is the head of the word, it is not uncommon (given the Asymmetry-Complexity Principle) for such heads to allow greater complexity and/or more marked structures than dependents in the same domain. Thus, assuming that for the parameter WBP the value 'yes' (i.e. codas are weightful) is the marked option, we are not surprised to find that in head syllables, the value of the parameter is yes (marked), while in dependent syllables it is no (unmarked).

If head-dependent asymmetries are a general and pervasive characteristic of languages, it should not greatly surprise us to find that the setting of parameters can be sensitive to the distinction between head and dependent positions. In fact, this is precisely what we expect, given that head/dependent asymmetries are generally present, and given that, in general, heads show greater complexity or markedness. Thus, we propose then that parameter settings *can*, if not typically *are*, relative to head positions, setting the marked value for the head.

A small point is that we do not, in fact, have to specify both values if we adopt, as a general convention, that the activation of the head option in (42c) implies that elsewhere the unmarked value obtains:

| | | | |
|---|---|---|---|
| (42) | HDP: | a. | Foot (left-headed) |
| | (revised) | b. | Word (left-headed) |
| | | c. | WBP-head (yes) |

We stress that this kind of context-sensitivity is far from arbitrary and this revision of the PPT (if indeed this is a revision rather than a clarification) is by no means an ad hoc move. We have, in fact, improved the theory by generalizing the impact of the *asymmetrical* nature of representations to the parameters that co-characterize these representations.

## 4.3 Explaining positional faithfulness

Appealing to differences between head and non-head positions has become rather common in OT. Implicit reference to the head – non-head distinction is seen, for example, in Beckman (1998) as an instantiation of *Positional Faithfulness*. Beckman states that there "is a small inventory of privileged linguistic positions which play a central role in the phonology of the world's languages" and she gives the following examples:

(43)

| | a. *Privileged positions* | b. *Non-privileged positions* |
|---|---|---|
| | -root-initial syllables | -non-initial syllables |
| | -stressed syllables | -unstressed syllables |
| | -syllable onsets | -syllable codas |
| | -roots | -affixes, clitics, function words |
| | -long vowels | -short vowels |

However, since proponents of OT make no commitments to any carefully formulated theory of representations, no attempt is made to explain *why* certain positions are more privileged. Noting this failure, Dresher and van der Hulst (1998) claim that the explanation lies in the fact that heads allow greater complexity (greater markedness) than dependents, and they suggest that an effort be made to analyze all instances of privileged positions as head positions. We believe that this is an interesting way to go, more interesting for sure than appealing to a notion of 'privilege' or 'salience' in the context of (unformulated) theories about production, perception and acquisition.

In the context of this article, we can only make suggestions as to how we can analyze the above-mentioned privileged positions as heads:

(44) a. *Privileged positions as heads*

| | |
|---|---|
| -root-initial syllables | Domain edge syllables are typical heads |
| -stressed syllables | Stressed syllables are heads |
| -syllable onsets | Onset consonants are heads (of onsets) and are heads of preceding codas in Coda Licensing |

| | |
|---|---|
| -roots | Roots typically form the head foot of prosodic words |
| -long vowels | Long vowels are headed domains (cf. Section 3.3.)[66] |

The greater faithfulness of 'privileged' positions is, in our view, due to the fact that heads allow the maximum complexity allowed by the grammar, which is formally accounted for by having the marked setting of the relevant parameter. Even though the faithfulness/complexity asymmetries can be expressed in an OT-analysis, we claim that our approach accounts for them in a principled manner that relies on the most fundamental structural assumption of our theory: the asymmetrical head-dependency relation.

## 5. Conclusions

In this article, we have motivated, explained, exemplified and improved a two-level, declarative constraint-based approach to phonology. As inevitable in this day and age, we have compared this approach to that of Optimality Theory, taking issue with the mechanism of ranking, and, more implicitly with the lack of commitment to any specific theory of representations.

A second important objective was to discuss the central role of head-dependency relations in phonology. Following the lead of Dependency and Government Phonology, we have advocated that HD-relations form the foundation of phonological representations at all layers (segmental, syllabic, and so on).

Even though claims have been made about the post-lexical system, our focus has been on the lexical system. Our main argument against OT is directed at the treatment of lexical phonology, since we consistently find more restrictive analyses that do not invoke superficial, construction-specific constraints and parochial constraint ranking. A logical possibility is that whereas the lexical system is best handled within the more restricted approach of the principles-and-parameters model that we advocate here, the post-lexical system (that is obviously closer to the everlasting battle between ease of articulation and ease of perception) is best handled by the much more powerful OT-style model. It is perhaps not a coincidence that, in certain quarters, OT-phonology has moved into the treatment of low-level phonetic detail, operating with constraints that refer to gradient properties involving formant frequencies and the like (cf. Boersma 1998).[67]

## Notes

**1.** The derivationality, in this case, forms part of the architecture of the grammar, rather than resulting from ad-hoc rule ordering.

**2.** This article pursues some of the points in van der Hulst and Ritter (2000b) regarding these same issues.

**3.** The proposition *[+nasal, +voice] is logically equivalent to propositions [+nasal] -> [-voice], and [+voice] -> [-nasal].

**4.** Shibatani (1973), Sommerstein (1974) and Clayton (1976) offer early proposals to use 'output constraints'. In general, all proposals to abandon extrinsic ordering, going back to Koutsoudas, Sanders and Noll (1974); Vennemann (1984); Ringen (1977); and Hooper (1976) lead to non-procedural phonologies. Strictly declarative frameworks are offered in Scobbie (1997); Bird (1990, 1995); and Coleman (1992, 1995a, b). Goldsmith (1993); Lakoff (1993); Karttunen (1993) and Orgun (1995), among others, present constraint-based systems that use more than one level. Prince and Smolensky (1993) also mention a long list of researchers that have anticipated their OT constraint-based approach.

**5.** cf. Trubetzkoy (2001:11): "The description of a phonological system of a language comprises *word phonology*, divided into lexical and morphological phonology, and *sentence phonology*."

**6.** Within lexical phonology, a distinction is usually made between two or more lexical levels, which correspond to blocks of morphological rules. We assume that the lexical constraint system is sensitive to a distinction between analytic and non-analytic domains (as proposed in Kaye 1995), a proposal that we discuss below.

**7.** Of course, our view of the lexicon here is simplified to the extent that the lexicon is seen as the collection of stored words. Obviously, both smaller (morphemes) and larger (phrases) units than words are stored. Cf. Jackendoff (2002) for extensive discussion.

**8.** Technically, this means that there is a system that produces the set of all possible combinations, a 'generator' of some kind. This system, of course, has no interesting linguistic properties or 'psychological reality' as it stands. As usual, viewing grammar as a checking system forces one to remain neutral with respect to the 'psychological' question as to what drives people to utter a certain combination of morphemes and words. The standard syntactic-centric view is not meant to be anymore 'psychologically 'real'. The pursuit of psychological reality drives Seuren (2001) to take the semantic structure as 'basic'. Perhaps, however, the pursuit of psychological reality should focus on the 'generator' and specify how speakers select combinations that they wish to submit to evaluation by the three checking systems.

**9.** In our view, there are no well-formedness requirements on morphemes as such. In reality (due to 'Lexicon Optimization'; cf. Prince & Smolensky 1993), morphemes tend to be largely well-formed. However, they can be potentially incomplete, either phonologically (or semantically). We, therefore, do not envisage a set of phonological 'morpheme structure constraints'; cf. Paradis & Prunet (1993) and Booij (1999) for discussion.

**10.** Later we suggest that the set of phonological expressions may not be infinite, because there is no recursion.

**11.** Thus, we take allophonic in a broad sense of automatic alternations, which could be neutralizing or not.

**12.** In the model that we follow there are no features. Rather 'phonemes' are truly indivisible **elements** that can occur alone or in combination to form 'segments' (i.e. units that occupy syllabic positions); cf. van der Hulst and Ritter (in prep.) for discussion. This idea is taken from such approaches as Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990) and Dependency Phonology (Anderson & Ewen 1987). The universal set of elements is so small, that all languages use all of them. Languages differ only in their constraints on element combinations (called *licensing constraints* in the GP literature); cf. Ritter 1995, 1999 for discussion of this topic.

**13.** In Section 3.3, we will comment on the meaning of the term 'phonetic interpretation' in this context.

**14.** So far, we seem to claim that all phonological alternation involves some sort of 'deletion', here understood as the failure of an element to be phonetically interpreted. In Section 3.3, we show how the role of 'phonetic interpretation' can also be held responsible for phenomena that, traditionally, call for 'insertion'.

**15.** Such crazy morphologized rules have been called *morpho-lexical rules* (in Anderson 1976), and *allomorphy rules* (in Aronoff 1976) and *distributional rules* (in Hudson 1974 and Hooper 1976). Cf. Dressler (1985) for a broad discussion of relevant phenomena.

**16.** We will avoid the term 'prosodic word' since we wish to distinguish between a phonological notion of word at the lexical and the post-lexical level, which we will refer to as phonological L-words and phonological PL-words, respectively.

**17.** The term 'phoneme' for segment-sized units turned inadequate when phonologists started arguing that phonemes consist of smaller parts, i.e. features. In our approach, the smallest phonological units are elements; this is why we say that these elements could be regarded as the true phonemes. This being said, we do not use the term 'phoneme' in order to avoid confusion due to its traditional usage.

**18.** This is not just a terminological point, involving replacing 'syllable' by 'package'. The relationship between onsets and rhymes is not regarded as establishing a constituent. Cf. van der Hulst and Ritter (1999) for discussion.

**19.** Inter-nuclear relationships are discussed in Section 3.3. Van der Hulst (to appear a) and van der Hulst & Rowicka (1997) argue that these relationships have a 'foot-like' appearance. We will not pursue this terminology here, however, pending a more thorough discussion of the notion foot in the L-phonology.

**20.** The distinction between 'inner' and 'outer' does not necessarily correlate with unproductive and productive word formation, although there certainly seems to be a tendency for the latter class to be highly productive.

**21.** Since many of her examples involve non-neutralizing alternations, the idea that lexical rules must be 'structure preserving' (as originally assumed by Kiparsky 1982, 1985) cannot be maintained.

**22.** For similar reasons, Borowsky assumes that level II suffixes go through the word-level phonology.

**23.** Cf. Brockhaus (1995) for an analysis of final devoicing in German within the GP framework.

**24.** In Kiparsky (1985) rules of this type are acknowledged as 'post-cyclic lexical rules', also referred to as 'word level' rules. Clearly these 'word-level' rules are distinct from the class of rules that Borowsky calls word-level rules. Cf. Booij and Rubach (199x) who specifically address the place of final devoicing in a lexical model.

**25.** It is, in addition, likely that the L-clitics, are subject to specific phonotactic constraints due to the fact that phonological constraints discriminate between strong (i.e. head) positions and weak (i.e. dependent) positions.

**26.** Repair (considered a potentially destructive type of operation) has also been explicitly barred from frameworks such as Declarative Phonology (Scobbie 1997; Bird 1990, 1995; Coleman 1995a, b). To account for apparent alternations in the form of morphemes and words, such models must rely rather heavily on underspecification or, in many cases, on denial of phonological generalizations and must appeal to lexical listing. Declarative Phonology, then, is very similar in spirit to GP.

**27.** In the model that we will discuss in Section 3.3, such rhymes are ruled out universally.

**28.** In Section 3.3, we discuss this example, including relevant data, in some detail.

**29.** The specific formulation of the constraint is not at issue here. As mentioned above, we would assume a constraint that specifies that the voice element is not licensed in the relevant context. Consequently it is not phonetically interpreted (rather than being deleted).

**30.** The original 'parse-and-fill' approach (Prince & Smolensky 1993), or the more recent proposals in Goldrick & Smolensky (1999) resemble our own views on 'repair' more closely than the latter versions of OT.

**31.** Apart from having to rank the 'anti-repair constraints', a second reason for ranking lies in replacing parameters by binary sets of constraints such that a parameter setting must be encoded by ranking the two 'choices'. We discuss this point in Section 4.

**32.** In practice, OT-analyses differ enormously from our type of analysis in terms of the kinds of well-formedness constraints that are used. Our approach, in fact, focuses on the form and substance of the actual constraints (principles and parameters). We address this issue briefly in Section 3.1.

**33.** The primitives of PL-phonology are, we assume, segments (skeletal positions plus all elements that are phonetically interpreted, i.e. licensed at the L-level). Thus, the PL-structure is built 'from scratch', and independent of the lexical structure, rather than taking up where the L-phonology has left off. PL-phonological words are aligned with L-words, irrespective of their L-syntactic (morphological) or L-phonological structure; cf. van der Hulst (to appear a) for discussion of these points.

**34.** Our notion of phonetic interpretation should not be confused with what is often called *phonetic implementation*. Since Pierrehumbert (1980), more attention has been paid to the relationship between phonological representations and their corresponding phonetic realizations. Pierrehumbert referred to the relationship as 'phonetic implementation', having in mind a system that converts a phonological representation into an actual phonetic event, an acoustic signal or articulatory movement(s). Subsequently, it has been proposed by some writers that post-lexical phonological processes *are* (part of) this implementation system. In our view, the post-lexical system is a *phonological* system that characterizes a set of well-formed representations. Phonetic implementation (i.e. the mapping of PL-phonological structure into a physical phonetic event) affects all phonological units that are licensed (i.e. phonetically interpreted) in this structure.

**35.** Cf. Anderson and Ewen (1987: 122): "...any sequences of word structures may be associated with several distinct utterance structures." The latter one-to-many relation is due to "information structure or considerations of rythm and tempo".

**36.** Anderson and Ewen (1987: 279) suggest that post-lexical structure may allow one head to have more than one dependent.

**37.** In van der Hulst and Ritter (2000a) we discuss a number of opacity effects within the framework of Head-driven Phonology, which we analyze without appeal to a lexical/post lexical distinction. Rather we show that opacity effects can also result from interpretation effects at the lexical level. More specifically, a rule in a traditional generative account that 'deletes' a unit that conditions another process will create a counterbleeding opacity effect. In our account, the unit that is said to be deleted would always still be there, but it would not be licensed and thus remain silent. The proposals in Goldrick & Smolensky (1999) go in a very similar direction.

**38.** Thus each constraint in Declarative Phonology expresses a 'true generalization' (cf. Hooper 1976). In general, it is believed that a ranking relationship between constraint pairs that stand in an 'elsewhere' (general - specific) relationship does not pose a threat to the declarative approach; cf. Scobbie (1997). Thus, our own approach allows the same kind of ranking.

**39.** It is usually said that opacity results from counterfeeding and counterbleeding relationships. Andrea Calabrese pointed out to us that feeding relations may also introduce opacity. If, in a string VCC, a vowel is inserted between the two consonants, the presence of the penultimate consonant might trigger a rule of intervocalic deletion. The resulting VVC string is opaque vis-à-vis the insertion rule, yet the ordering is feeding.

**40.** Cf. McCarthy (1999). This is, in a sense, ironic because at an abstract level, OT shares with SPE (Chomsky & Halle 1968) its need for extrinsic ordering, which in SPE was introduced precisely to deal with procedural opacity.

**41.** We use 'procedural' here in the way others use 'derivational' or 'serial' (cf. Clements 2000 for discussion).

**42.** OT uses an input - output distinction, but this is in itself not an automatic consequence of constraint ranking. Rather, this distinction is necessary in order to deal with repair in terms of anti-repair constraints.

**43.** Our own position in this article is no less programmatic of course.

44. Contrary to the position of Bromberger and Halle (1989), it is also assumed that phonology is not different from syntax in that neither have ordered rules. Cf. Ritter (1995) for a discussion of the role of general UG principles in both syntax and phonology.

45. The term 'level' is used in two different ways, here and in the literature at large. Sometimes it stands for the distinction between word and sentence phonology, while in other cases it refers to different 'layers' or 'domains' in the hierarchical organization of representations.

46. Following Jackendoff (2002), we extend this point, and all following points about the parallelism between grammatical systems, to the semantic system.

47. It follows, we think, that adopting a non-OT approach to syntax, entails a similar stance for phonology, and vice versa.

48. It is sometimes remarked that the notion head is differently used in phonology and syntax. This may be so. We believe that there are also deep and significant resemblances. The 'definition' of what constitutes a head may be too narrow and too domain-specific on both sides, and it may be too early to come to a full understanding of the relevant concept that is shared by both modules.

49. New names are always dangerous. We do not wish to imply here that our work is the phonological counterpart of what is known as 'Head-Driven Phrase Structure Grammar'(HPSG). Such a parallelism would require embracing the specific approach of Declarative Phonology and combining it with the structural ideas of Government and Dependency Phonology. This, of course, is possible; cf. Coleman (1992). However, on the other hand it is probably true that HPSG and HDP share a concern with on-procedural grammars in which head - dependent relations play a pivotal role.

50. At the level of foot structure, it has been proposed that all feet are binary. In most theories, however, this binarity may involve a non-branching foot if the only syllable is itself branching (i.e. bimoraic, however moraicity is expressed).

51. In Government phonology, and in HDP, the relationship between the onset and the rhyme is not seen as a constituent relationship. The 'syllable' is not understood as a constituent in the traditional sense. Given this, the relationship in (22b) between onset and rhyme is not one of constituency, although it is a head-dependent relationship. GP proponents usually do not abstract away from the linear order of onset and rhyme. Hence, the HD-relationship in that case is right-headed.

52. As Dresher and van der Hulst (1998) point out, this principle is relevant for αα–relationships only.

53. The division into constituent, interconstituent and edge constraints bears an abstract relationship to the typology of rules proposed in Selkirk (1980) that involves: domain-span rules, juncture rules and edge rules.

54. Neither HDP nor Government Phonology are, at this point, restricted enough to lead to unique analyses. There are alternatives and the discussion is ongoing, just like in any other framework that we are aware of. For discussion of different analyses of the Yawelmani data, cf. Rowicka (1999).

**55.** There may be crosslinguistic variation in the phonetic properties of the realization of ungoverned empty nuclei. This may suggest a parameterized aspect to interpretation.

**56.** In this section, following Kaye, Lowenstamm and Vergnaud (1990), we propose an analysis that uses a right-headed relationship. Van der Hulst and Rowicka (1997) propose a left-headed relationship. For a discussion of the two analyses, cf. Rowicka (1999). In this article, we do not wish to get involved in a discussion of various alternative analyses that are consistent with the basic ideas of Government Phonology.

**57.** A complete discussion of the theory would include a more subtle discussion of the circumstances that determine the choice between 'silent' and 'audible' empty nuclei.

**58.** Even though the suffix vowel is /i/, we do not assume that it corresponds to an empty nucleus; there is no alternation to warrant such an analysis. But even if we assumed it was empty, it would have to be realized because the suffixal nucleus would have to be audible because it does not meet the conditions in (29).

**59.** This fact does not necessitate the insertion of a phonological element.

**60.** At this point, we include the X-positions (which we assumed all along) in our notation.

**61.** It is unfortunately the case that there has been a tendency, especially in OT, to think of constraints as specifying very concrete, almost 'phonetic' properties of words and utterances. Our constraints typically make reference to more abstract phonological properties. This difference (to the extent that it exists), however, has no bearing on the use of the term constraint. Even in OT there is no need for constraints to be literally superficial and, indeed, the more interesting constraints are not superficial, but simply mimic traditional principles and parameters.

**62.** This claim is notoriously unfalsifiable, since constraints need not have an effect in languages in which their demands are totally superceded by higher-ranked constraints. It is also implausible in view of the ad hoc constraints that appear in the OT-literature.

**63.** We are ignoring here that OT has substantially enriched the theory of constraint interaction with all sorts of additional mechanisms, such as *extrinsic ordering, universally top-ranked constraints* and more recently*: constraint conjunction, sympathy theory et alia.* This further weakens the claim of being 'conceptually simple'.

**64.** In their actual analysis, Rosenthall & van der Hulst show that WBP itself interacts with a constraint '*Append' which rules out the adjunction of a non-moraic coda consonant to the syllable node. We suppress that complication here. Note, by the way, that NotAppend is virtually the opposite of No-WBP.

**65.** The phonetic evidence for taking the leftmost stress as primary is controversial, but we see no reason to demand, in general, that primary stresses are phonetically more salient than secondary stresses. The notion 'stress' refers to the phonetic exponents of a metrical structure. General principles of well-formedness demand that all structures must have a head, and that this head must be on the edge of the constituent domain.

**66.** Short vowels are heads of their nucleus/rhyme, but long vowels govern a sister; this is the relevant difference.

## References

Allen, M. (1978). Morphological investigations. Doctoral dissertation, University of Connecticut.

Anderson, S. R. (1976). On the interaction of phonological rules of various types. *Journal of Linguistics, 11*, 39–62.

Anderson, J. & Ewen. C. (1987). *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

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

Beckman, J. (1998). Positional faithfulness. Doctoral dissertation, University of Massachusetts, Amherst.

Bird, S. (1990). Constraint-based phonology. Doctoral dissertation, University of Edinburgh.

Bird, S. (1995). *Computational Phonology. A Constraint-Based Approach*. Cambridge: Cambridge University Press.

Boersma, P. (1998). Functional phonology: formalising the interactions between articulatory and perceptual drives. Doctoral dissertation, University of Amsterdam.

Booij, G. (1999). Morpheme structure constraints and the phonotactics of Dutch. In H. van der Hulst and N. A. Ritter (Eds.), *The Syllable: Views and Facts* (pp. 53–68). Berlin: Mouton de Gruyter.

Booij, G. & Rubach, J. (1987). Postcyclic versus postlexical rules in lexical phonology. *Linguistic Inquiry, 18*, 1–44.

Borowsky, T. (1993). On the word level. In E. Kaisse and S. Hargus (Eds.), *Studies in Lexical Phonology. Phonetics and Phonology. Volume 4* (pp. 199–234). New York: Academic Press.

Brockhaus, W. (1995). *Final Devoicing in the Phonology of German*. Tübingen: Max Niemeyer Verlag.

Bromberger, S. & Halle, M. (1989). Why phonology is different. *Linguistic Inquiry, 20*, 51–70.

Calabrese, A. (1986). The interaction of phonological rules and filters in Salentino. *NELS, XVII*, 79–88.

Calabrese, A. (1995). A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry, 26*, 373–463.

Calabrese, A. (To appear). Dynamic phonology. MS, University of Connecticut.

Charette, M. (1990). License to govern. *Phonology, 7*, 233–253.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.

Clayton, M. (1976). The redundancy of morpheme structure conditions. *Language, 52*, 296–313.

Clements, G. N. (2000). In defense of serialism. In N. A. Ritter (Ed.), *A Review of Optimality Theory* (pp. 181–190). [Special Triple Issue of *The Linguistic Review, 17*, 2–4].

Coleman, J. (1992). The phonetic interpretation of headed phonological structures containing overlapping constituents. *Phonology, 9*, 1–44.

Coleman, J. (1995a). Declarative lexical phonology. In J. Durand and Katamba, F. (Eds.), *Frontiers in Phonology: Atoms, Structures, Derivations* (pp. 333–382). Essex: Longman.

Coleman. J. (1995b). *Phonological Representations. Their Names, Forms, and Powers*. Cambridge: Cambridge University Press.

Dell, F. & Selkirk, L. (1978). On a morphologically governed vowel alternation in French. In S. J. Keyser (Ed.), *Recent Transformational Studies in European Languages* (pp. 1–52). Cambridge, MA: MIT Press.

Dresher, E. & van der Hulst, H. (1998). Head-dependency in phonology: complexity and visibility. *Phonology, 15*, 317–352.

Dressler, W. (1985). *The Dynamics of Derivation*. Ann Arbor: Karoma Publishers, Inc.

Fudge, E. (1969). Syllables. *Journal of Linguistics, 5*, 253–287.

Fudge, E. (1987). Branching structure within the syllable. *Journal of Linguistics, 23*, 359–377.

Goldrick, M. & Smolensky, P. (1999). Opacity and turbid representations in Optimality Theory. Talk presented at the Chicago Linguistic Society, April 22, (1999).

Goldsmith, J. (1993). Harmonic phonology. In Goldsmith, J. (Ed.), *The Last Phonological Rule* (pp. 21–60). Chicago and London: The University of Chicago Press.

Gussman, E. & Kaye, J. (1993). Polish notes from a Dubrovnik café. I: the yers. *SOAS Working papers in linguistics and phonetics, 3*, 427–462.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry, 20*, 253–306.

Harris, J. (1990). Segmental complexity and phonological government. *Phonology, 7*, 255–300.

Hooper, J. (1976). *Introduction to Natural Generative Phonology*. New York: Academic Press.

Hudson, G. (1974). The representation of non-productive alternations. In J. Anderson and C. Jones (Eds.), *Proceedings of the First International Conference on Historical Linguistics* [Volume II] (pp. 203–229). Amsterdam: North Holland.

Hulst, H. van der (1984). *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.

Hulst, H. van der (2000). Modularity and modality in phonology. In N. Burton-Roberts, P. Carr and Docherty, G. (Eds.), *Phonological Knowledge: Its Nature and Status* (pp. 207–244). Oxford: Oxford University Press.

Hulst, H. van der (To appear a). Structure paradoxes in phonology. In *Festschrift*.

Hulst, H. van der (To appear b). Stress and accent. To appear in *Encyclopedia of Cognitive Science*. MacMillan Reference Ltd.

Hulst, H. van der & Ritter, N. A. (1999). Head-driven phonology. In H. van der Hulst and N. A. Ritter (Eds.), *The Syllable: Views and Facts* (pp. 113–168). Berlin: Mouton de Gruyter.

Hulst, H. van der & Ritter, N. A. (2000a). No sympathy for opacity. In S. Billings and A. Griffith (Eds.), *The Main Session [CLS 35]* (pp. 153–173). Chicago: Chicago Linguistics Society.

Hulst, H. van der & Ritter, N. A. (2000b). The SPE-heritage of Optimality Theory. In N. A. Ritter (Ed.), *A Review of Optimality Theory* (pp. 259–290). [Special Triple Issue of *The Linguistic Review, 17* 2000, 2–4].

Hulst, H. van der & Ritter, N. A. (In preparation). *The Syntax of Segments*. Monograph, HIL/Leiden University.

Hulst, H. van der & Rowicka, G. (1997). On some parallels between unrealized empty nuclei and unstressed syllables. In G. Booij and J. van de Weijer (Eds.), *HIL phonology papers* III [HIL publications #5] (pp. 125–149). The Hague: Holland Academic Graphics.

Inkelas, S. (1989). Prosodic constituency in the lexicon. Doctoral dissertation, Stanford University.

Jackendoff, R. (2002). *Foundations of Language*. Oxford: Oxford University Press.
Karttunen, L. (1993). Finite-state constraints. In Goldsmith, J. (Ed.), *The Last Phonological Rule. Reflections on constraints and derivations.* (pp. 173–194). Chicago and London: The University of Chicago Press.
Kaye, J. (1989). *Phonology: A Cognitive View.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.
Kaye, J. (1990). 'Coda' Licensing. *Phonology, 7*, 301–330.
Kaye, J. (1995). Derivations and interfaces. In J. Durand, J. and Katamba, F. (Eds.), *Frontiers in Phonology: Atoms, Structures, Derivations* (pp. 289–332). Essex: Longman.
Kaye, J., Lowenstamm, J. & Vergnaud, J.-R. (1985). The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook, 2*, 305–328.
Kaye, J., Lowenstamm, J. & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology, 7*. 193–232.
Kenstowicz, M. & Kisseberth, C. (1979). *Generative Phonology. Description and Theory*. New York: Academic Press.
Kiparsky, P. (1982). From cyclic phonology to lexical phonology. In H. van der Hulst and N. Smith (Eds.), *The Structure of Phonological Representations* [Part I] (pp. 131–177). Dordrecht: Foris.
Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology Yearbook, 2*, 85–139.
Kiparsky, P. (2000). Opacity and cyclicity. In Ritter, N. A. (Ed.), *A Review of Optimality Theory* (pp 351–366). [Special Triple Issue of *The Linguistic Review, 17*, 2–4].
Kisseberth, C. (1970). On the functional unity of phonological rules. *Linguistic Inquiry, 1*, 291–306.
Koutsoudas, A., Sanders, G. & Noll, G. (1974). On the application of phonological rules. *Language, 50*, 1–28.
Lakoff, G. (1993). Cognitive phonology. In J. Goldsmith (Ed.), *The Last Phonological Rule* (pp. 117–145). Chicago and London: The University of Chicago Press.
McCarthy, J. (1995). Emergence of the unmarked. Ms., University of Massachusetts, Amherst.
McCarthy, J. (1999). Sympathy and phonological opacity. *Phonology, 16*, 331–399.
Mohanan, K. P. (2000). The theoretical substance of the optimality formalism. In N. A. Ritter (Ed.), *A Review of Optimality Theory* (pp. 143–166.) [Special Triple Issue of *The Linguistic Review, 17*/2–4].
Orgun, O. (1995). A declarative theory of phonology-morphology interleaving. Hand out, University of California, Berkeley.
Paradis, C. (1988). On constraints and repair strategies. *The Linguistic Review, 6*, 71–97.
Paradis, C. & Prunet, J.-P. (1993). On the validity of morpheme structure constraints. In C. Paradis and D. LaCharité (Eds.), *Constraint-based Theories in Multilinear Phonology* (pp. 235–256). [Special issue of *The Canadian Journal of Linguistics, 38*].
Pierrehumbert, J. (1980). The phonetics and phonology of English intonation. Doctoral dissertation, MIT.
Prince, A. & Smolensky, P. (1993). Optimality Theory and constraint interaction in generative grammar. Technical Report #2 of the Rutgers Center for Cognitive Science. Rutgers University, New Jersey.

Ringen, C. (1977). Vowel harmony: theoretical implications. Doctoral dissertation, Indiana University.

Ritter, N.A. (1995). The role of Universal Grammar in phonology: a government phonology approach to Hungarian. Doctoral dissertation, New York University.

Rosenthall, S. & H. van der Hulst (1999). Weight by position by position. *Natural Language and Linguistic Theory, 17* (3), 499–540.

Rowicka, G. (1999). *Ghost vowels. A strict CV approach* [LOT/HIL Dissertation # 16]. The Hague: HAG.

Scobbie, J. (1997). *Autosegmental Representation in a Declarative Constraint-based Framework*. New York & London: Garland press.

Selkirk, L. (1980). Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff and M. Kean (Eds.), *Juncture*. (pp. 107–129). Saratoga: C.A. Anma Libri.

Seuren, P. (2001). *A View of Language*. Oxford: Oxford University Press.

Siegel, D. (1984). Topics in English morphology. Doctoral dissertation, MIT.

Shibatani, M. (1973). The role of surface phonetic constraints in generative phonology. *Language, 49*, 87–106.

Singh, R. (1987). Well-formedness conditions and phonological theory. In W. Dressler (Ed.), *Phonologica (1984)* (pp. 273–286). Cambridge: Cambridge University Press.

Sommerstein, A. H. (1974). On phonotactically-motivated rules. *Journal of Linguistics, 10*, 71–94.

Steriade, D. (2000). P-map effects on constraint organization. Talk presented at the First North American Phonology Conference, Montreal May 27, 2000.

Stewart, J. (1983). Akan vowel harmony: The word structure conditions and the floating vowels. *Studies in African linguistics, 14*.

Strauss, S. (1982). *Lexicalist Phonology of English and German*. Dordrecht: Foris Publications.

Trubetzkoy, N. S. (2001). *Studies in General Linguistics and Language Structure*. Edited by A. Liberman. Durham & London: Duke University Press.

Vennemann, T. (1984). Words and syllables in natural generative grammar. In A. Bruck, R. A. Fox and M. W. La Galy (Eds.), *Papers from the Parasession on Natural Phonology* (pp. 346–374). Chicago: Chicago Linguistic Society.

Yoshida, S. (1993). The case of Palestinian vowel harmony. *The Linguistic Review, 10*, 127–159.

# Obstruent neutrality in nasal harmony*

G. L. Piggott
McGill University

## 1. Introduction

In long distance nasalization or nasal harmony, segments are either targets of the process or remain neutral. Cross-linguistically, vowels and laryngeals are consistently targets, while sonorant consonants (i.e., glides, liquids) vary between neutral and target status. In contrast, the behaviour of obstruents is less uniform. While fricatives may also be either neutral or target segments, stops as a class may be neutral but are never targets. Among stops, the voiced subset may be targeted but the voiceless set is resolutely neutral. When voiced stops are targeted, other voiced segments must be also. The neutrality of sonorants and fricatives takes the form of opacity to nasal harmony. Neutral stops, by comparison, are either opaque as a class, or the voiceless subset is transparent. When the latter situation arises, voiceless fricatives are (generally) also transparent. Consequently, the robust generalization is that transparency is restricted to the class of voiceless obstruents. No voiced consonant (sonorant or obstruent) is ever transparent to nasal harmony. The following generalizations about segment neutrality must, therefore, be explained by any explanatory analysis of nasal harmony.

(1) Segment neutrality

a. Sonorant consonants and fricatives may be either targets or neutral segments.
b. Neutral sonorants are always opaque to nasal harmony.
c. The class of obstruent stops is never targeted.
d. Voiceless stops are always neutral.
e. If voiceless stops are transparent, then voiceless fricatives must be also.
f. Only voiceless obstruents are transparent to nasal harmony.

These generalizations reveal an asymmetry between the behaviour of obstruent stops and other segment types; nasal harmony may target any segment class except stops.

Among the most recent attempts to capture the above generalizations, analyses of nasal harmony by Piggott (1992), Piggott and van der Hulst (1997) and Walker (1998) stand out as the most comprehensive. Piggott's (1992) analysis is in a derivational framework where feature-geometric configurations play crucial roles. Its main weakness is in the absence of justification for some of the representations, independent of the facts of nasal harmony. It also fails to explain why voiced obstruents are never transparent to nasal harmony (1f). The Piggott and van der Hulst (1997) account relies less on arbitrary feature-geometric representations, but it does not address the restriction in (1d) and also fails to explain (1f). Walker (1998) proposes an alternative analysis in which neither serial derivation nor feature-geometry plays any role. However, this analysis suffers from some of the same failures as its predecessors. Crucially, it does not explain the universal neutrality of voiceless stops (1d), the dependency between stop and fricative transparency (1e) and the absence of transparent voiced obstruents (1d). The present paper explains why segment neutrality in nasal harmony is manifested as summarized in (1), with special attention to the behaviour of obstruents. The novel feature of this new analysis is the link it establishes between neutrality and the structure of consonant inventories.

The theoretical framework of this paper is also somewhat novel. While it assumes that phonological patterns must satisfy a set of constraints, the approach to constraint interaction differs from the one endorsed by Optimality Theory (OT) (Prince & Smolensky 1993) which allows for the random ordering of constraints. In fact, I argue that a theory that adheres to the fundamental tenets of OT is, in principle, incapable of explaining the behaviour of obstruent stops in nasal harmony patterns. The proposed alternative to standard OT is a framework in which constraints have the status of either principles or parameters of variation, in the spirit of Chomsky (1981).

The outline of the rest of this paper is as follows. Section 1.1 provides a brief overview of the crucial differences between two views of constraints and constraint-interaction; the standard OT is contrasted with the principles and parameters (P&P) framework assumed in this paper. It indicates how constraints may be ranked, although ranking is non-random. Section 2 presents illustrative examples of the various attested patterns of nasal harmony. The various subsections of 3 develop an analysis that explains the neutrality of stops and fricatives, while also allowing for the systematic targeting of fricatives. Section 4 provides a brief explanation for how and why fricatives, liquids and glides vary between neutral and target status. Finally, Section 5 evaluates why a constraint-based approach like standard OT cannot explain all of the generalizations in (1), especially the behaviour of obstruent stops.

### 1.1 Constraints and constraint-interaction

Since the early 70's, the generative phonology program has embraced the idea that surface phonological forms (i.e. outputs) often have to satisfy some well-formedness conditions (i.e. constraints) (cf. Kisseberth 1970).[1] OT has given new impetus to this program by making constraint satisfaction the central issue in linguistic analysis. The surface representations or outputs in a particular language are the candidates that are evaluated as best satisfying the demands of a set of constraints. In the OT framework, the set is the same for all languages, but languages may arbitrarily impose different rankings on these constraints. Arguments for language-specific ranking of constraints are based on cases where two constraints are demonstrably active in a language but make conflicting demands. In such instances, the conflict can be resolved by allowing the demands of one to supersede the other. However, the empirical evidence from many of these cases merely show that a more specific constraint must have precedence over a more general one. Therefore, constraint ranking can often be attributed to a principle like the *Elsewhere Condition* proposed by Kiparsky (1973).[2] This condition is considered to be one expression of a theorem attributed to the Sanskrit grammarian, Pāṇini. The following adaptation readily allows for the extension of this theorem to the interaction of constraints.

(2) Pāṇini's Theorem (PT)
Given a pair of constraints **A** and **B**, where **B** has scope over the set of elements X and **A** has scope over the set of elements Y where Y is a proper subset of X, **A** (the special case) takes precedence over **B** (the general case), if the demands of **A and B** are mutually contradictory in some domain **D** (**i.e. special » general**).

A situation of mutual contradiction is deemed to exist if the satisfaction of **A** might be impeded by the satisfaction of **B** and vice-versa. In the above formalization of PT, the subset relation between the elements targeted by **A** and **B** is crucial. Whenever a language provides empirical evidence for the presence of such constraints, their ranking is predetermined; **A** must precede **B**. In other words, PT makes the testable claim that direct evidence for the ranking **B** » **A** could not be available.[3]

*Pāṇini's Theorem* is adopted as a meta-condition on constraint ranking in the framework of this paper. I also adopt another meta-condition requiring constraints regulating structure to be superordinate to the so-called FAITH constraints that regulate featural or segmental preservation (cf. Paradis 1988; Piggott 1997). For example, a demand that underlying segments must be preserved in outputs cannot be satisfied at the expense of a syllable well-formedness condition dictating what are possible coda segments. In addition to such conditions on ranking, the framework of this paper embraces a proposal by Prince and Smolensky (1993) that attributes some crucial rankings of constraints to the demands of some fixed harmonic scale. Constraints are considered to belong to one of two types. One type is imposed by Universal Grammar (UG) on all languages and has the status of principles. The other type consists of constraints that are activated on a language-particular basis; they belong to the parameters of variation. Grammatical differences can therefore be correlated with the choice to activate or not activate particular constraints. The conception of constraints as principles or parameters is one of the ways in which the theoretical framework of this paper differs from standard optimality-theoretic thinking. However, it incorporates some of the OT ideas such as the principle of parallelism, requiring that a set of potential surface forms be evaluated together to determine which one best satisfies the requirements of the set of active constraints. The following is a summary of the basic tenets of the principles and parameters (P&P) framework.

(3) Fundamentals of P&P framework
- a. A universal constraint set of two types
  - i. Principles (regulating universal properties)
  - ii. Parameters (language-particular choices)
- b. A freely generated candidate set of outputs
- c. All precedence relations between constraints are universally determined
- d. Parallel evaluation of candidate outputs by a constraint set.

Constraints that are not ranked in accordance with some fixed scale or with the demands of some meta-condition would remain unranked.

In a theoretical framework that requires constraint ranking to be universally determined, it is interesting and informative to see how a pair of antagonistic but unranked constraints would fare. To illustrate this point we may consider a hypothetical constraint **C**, requiring that some input feature (**P**) be present in outputs, and a second constraint ***C**, requiring that **P** not be present in outputs. The candidate set to be evaluated by these constraints logically includes those that contain **P** and those that lack P. Assuming that candidates are freely generated, there must be a competing pair ($Cand_p$, $Cand_{\sim p}$) where the difference between them is the presence versus absence of P.

(4)

| | Input: P | C | *C |
|---|---|---|---|
| ☞ | a. $Cand_p$ | | *! |
| ☞ | b. $Cand_{\sim p}$) | *! | |

Since each of the competitors incurs a violation of one of the antogonistic constraints, the evaluation metric cannot select one over the other. Therefore, when antagonistic constraints are unranked, the resulting grammar would make the logically incoherent claim that two completely contradictory statements are both correct. To avoid such a contradiction, phonological theory must require the evaluation metric to designate a single output as optimal. The P&P framework of this paper assumes the restriction that rules out the possibility of (4). Consequently, if the hypothetical constraints **C** is a principle, then the theory automatically excludes ***C** from the universal constraint set. If, on the other hand, **C** and ***C** are both parameters of variation, then neither is automatically excluded from the universal set, but only one of these may be active in a particular grammar. In effect, a grammar with a constraint interaction like that in (4) is disallowed.[4]

## 2. Nasal harmony patterns

Various surveys of nasal harmony (e.g. Cohn 1987; Piggott 1992; Walker 1998) identify four patterns in which some segment blocks the extension of nasality. Examples of languages illustrating these patterns are provided in (5).

(5) a. Sundanese (Robins 1957)

| | | | |
|---|---|---|---|
| ɲãĩ̃ãn | 'wet' | mõlohok | 'stare' |
| mĩʔãsih | 'love' | ŋãtur | 'arrange' |
| mãwur | 'spread' | ŋũdag | 'pursue' |

b. Malay (Onn 1976)

| | | | |
|---|---|---|---|
| mãỹãn | 'stalk' | mə̃laran | 'forbid' |
| mẽw̃ãh̃ | 'be luxurious' | mãkan | 'eat' |
| mãʔãp | 'pardon' | | |

c. Kolokuma Ijo (Williamson 1965)

| | | | |
|---|---|---|---|
| bẽĩ̃(n)[5] | 'be full' | kɔ̃ɾ̃ɔ̃ŋgbɔː | 'thin' |

| | | | |
|---|---|---|---|
| õw̃ĩ(n) | 'bite' | izõŋgo | 'jug' |
| ỹãr̃ĩ(n) | 'shake' | otõŋgbolo | 'mosquito' |

d. Applecross Gaelic (Ternes 1973)

| | | | |
|---|---|---|---|
| f̃r̃ãĩṽ | 'root' | str̃ãĩ:γ̃ | 'string' |
| k$^{h}$ɔ̃ĩs̃paxk | 'wasp' | mã:h̃ãr̃ | 'mother' |
| ʃ̃ɛ̃nɛ̃:ṽãr̃ | 'grandmother' | ãh̃ũç̃ | 'neck' |

In the first pattern (5a), vowels and laryngeals are targets, while glides (semivowels), liquids, fricatives and obstruent stops are blockers. The second pattern (5b) extends the targets to glides with a correlated reduction in the blocking set. In the third pattern (5c), the target set includes liquids, while the blockers are limited to the obstruents. Finally, the targets of nasalization in the fourth pattern (5d) include all segments except obstruent stops. Piggott (1992) categorizes these patterns as variants of one type called Type A. The variation is summarized in (6) where the shaded areas identify blocking segments. Because laryngeals are normally targeted, they are ignored in this summary.

(6) Targets and blockers in Type A harmony

| | Vowels | Glides | Liquids | Fricatives | Stops |
|---|---|---|---|---|---|
| A1 | | | | | |
| A2 | | | | | |
| A3 | | | | | |
| A4 | | | | | |

Segment in Type A harmony is subject to a number of entailments. For example, if glides are opaque to nasalization, then liquids and fricatives must be also. Note that obstruent stops are always among the blocking segments.

The profile of a Type A pattern is marked by the opacity of neutral segments. There is also a nasal harmony system, Type B, in which neutral segments are transparent. The latter occurs in Tucanoan languages like Southern Barasano (Smith & Smith 1971; Jones & Jones 1991) and Tuyuca (Barnes 1996), and in Tupi languages like Guarani (Gregores & Suarez 1967; Penner 1994). Type B nasal harmony is attested in only one pattern. The following data from Southern Barasano illustrate its salient features.

(7) Nasal harmony in Southern Barasano[6]

a. Nasal words

i.

| | |
|---|---|
| mãh̃ãɲɨ̃ | 'corner' |
| mãnõ | 'none' |
| ẽõnõ | 'mirror' |

b. Non-nasal words

| | |
|---|---|
| $^{m}$ba$^{ŋ}$go | 'eater' |
| ta$^{m}$boti | 'grass' |
| $^{n}$diro | 'fly' |

| | | | | |
|---|---|---|---|---|
| ii. | mãsã | 'people' | wesika | 'above' |
| | w̃ãtĩ | 'demon' | wati | 'going' |
| | kãmõkã | 'rattle' | hikori | 'tail' |
| iii. | mãhã-mã | 'go up!' | wa-$^{m}$ba | 'come!' |
| | mãsɨ̃-w̃ɨ̃ | 'I knew' | ʔoha-wɨ | 'I wrote' |
| | hũnẽ-nẽ | 'to hurt' | yi-re | 'to say' |
| | mĩnõ-ɲã | 'leaf stream' | $^{\text{ŋ}}$gahe-ya | 'another stream' |
| | yikõã-mɨ̃ | 'I did completely' | wa-$^{m}$bɨ | 'I went' |

The targets of nasalization in Southern Barasano are vowels, glides, liquids and voiced stops that are usually prenasalized. Voiced fricatives do not occur in this language, but this class of consonants is subject to nasalization in Guarani (e.g. nã-pẽ-pɨ̃tɨ̃ṽõ-ĩ 'he does not help you'). The unifying feature of the targeted segments in Type B harmony is therefore that they are voiced, while voiceless stops and fricatives are transparent to the extension of nasality.

(8) Targets and transparent segment in Type B harmony

| | Vowels | Glides | Liquids | Voiced fricatives | Voiced stops | V'less fricatives | V'less stops |
|---|---|---|---|---|---|---|---|
| B | | | | | | | |

Evidently, Type B languages manifest a surface contrast between a target group containing vowels, glides, liquids, voiced fricatives and voiced stops and a neutral group containing voiceless fricatives and voiceless stops. This contrast is not attested in any Type A language.

Both types of harmony are subject to restrictions on the direction of nasal extension. For example, the process is progressive in Sundanese and Southern Barasano but regressive in Kolokuma Ijo and Guarani. Obviously, each language must designate the direction of nasal extension from a trigger, which may be an underlying nasal consonant or vowel. When a nasal consonant is a trigger, its position in syllable structure appears to be relevant. Spreading from a cluster where a nasal in a coda position is followed by an onset consonant (i.e. NC clusters) is normally regressive. If an underlying NC cluster fuses into a simple nasal, nasality may spread progressively from the simplex nasal, as is the case in Bonggi (Boutin 2000) and other Austronesian languages. In contrast, spreading from a CN cluster is invariably progressive. The implication is that, in harmony systems, nasalization is never transmitted directly from one consonant to another. These observations are significant, because both NC and CN combinations may sometimes surface as NN clusters.[7]

## 3. The activation of nasal harmony

Before turning to the analysis of segment neutrality, we must first introduce the constraint that is best satisfied by requiring the long distance extension of the nasal feature. The process does not occur in every language. Spanish, German, French and English are among the languages where it is not attested. Therefore, the constraint that commands the extension of the nasality over sequences of segments is activated on a language-particular basis. While the exact formulation of this constraint is not a trivial matter, for the purpose of this paper, the following should be sufficient.

(9) NASAL HARMONY (NasHar)
A nasal feature must be extended over all segments in a word.[8]

Once activated by a particular language, the best satisfaction of this constraint requires that nasality be extended across a word until a non-nasalizable element is encountered or the edge of the word is reached. Elements that cannot be nasalized are expected to be opaque, because nasal harmony like other phonological relations must conform to a general locality condition that requires elements in a relation to be adjacent at the relevant structural level. This requirement, which has guided phonological analysis for the last two decades, is sometimes referred to as the no-gap condition (see Itô, Mester & Padgett 1995: 598).

Cross-linguistically, vowels readily bear the nasal feature, and we know from the patterns in (5) and (7) that, when nasal harmony occurs, vowels are always targets. It is therefore reasonable to infer that, if vowels cannot be nasalized, nasal harmony is impossible. Walker's (1998) analysis of nasal harmony attributes the blocking effect of vowels to a restriction on feature combinations that penalizes occurrences of nasal vowels. This idea can be readily adopted in the present analysis in the form of the following constraint.

(10) *NASAL-VOWEL (*NasV)
The surface representations must not contain nasal vowels.

Since the activation of (9) and (10) is a language-specific choice, there are, in principle, two ways to block nasal harmony. One option is to choose not to activate NasHar, while the other possibility results from the activation of *NasV. We allow for both possibilities, because, while a ban on nasal vowels would block nasal harmony, the absence of such a ban does not necessarily entail the possibility of harmony. For example, French clearly does not ban nasal vowels.

Nevertheless, nasal harmony must still be ruled out in this language, because nasality cannot be extended from any nasal segment.

The ban on nasal vowels can result in the blocking of nasal harmony only if *NasV invariably takes precedence over NasHar. This dominance does not have to be stipulated, however, because it follows directly from *Pāṇini's Theorem*. Within the domain of harmony, NasHar requires that all segments be nasalized, while *NasV excludes a subset of the segments (i.e. the vowels) from nasalization. Hence, the requirements of the latter, being more specific, are invariably satisfied at the cost of violating the more general NasHar. In other words, the ranking *NasV » NasHar is intrinsic to this pair of constraints and not imposed extrinsically by any arbitrary grammatical statement. The following tableau illustrates the consequences of this ranking in a hypothetical language that has activated both constraints

(11) Vowel opacity

| | Input: maʔa | *NasV | *NasHar |
|---|---|---|---|
| | a. mãʔa | *! | |
| | b. mãʔã | *!* | |
| ☞ | c. maʔa | | *** |

Although (11a, b) better satisfy the requirement that nasality be extended, each is condemned by the occurrence of nasal vowels.

### 3.1 The opacity of obstruent stops

While vowels are the most receptive to nasalization, obstruent stops are the most resistant. To account for the latter observation, we must assume that the satisfaction of some constraint may force stops to resist nasalization at the expense of demands of nasal harmony. Adapting ideas from McCarthy and Prince (1995) as extended by Causley (1999) and others, I postulate that the relevant constraint is a member of the max family.

(12) Max Stop (MaxStop)
An underlying stop must have an identical surface correspondent.

This constraint incorporates both segment correspondence and feature identity. It is fully satisfied only when underlying and surface correspondents have the same feature specification. Following a familiar idea in OT, I assume that each difference in feature is penalized. Therefore, violations of this constraint can be gradient.

There is obviously a potential conflict between MaxStop and NasHar, because the extension of nasality through an obstruent stop would prevent it from being a completely faithful correspondent of an underlying oral counterpart. On the other hand, if the identity of an oral obstruent stop is preserved in the domain of nasal harmony, nasal extension cannot be fully achieved. To ensure that the underlying features of obstruent stops are preserved in nasal harmony systems, MaxStop would have to be superordinate to NasHar. Such a ranking is exactly what PT decrees; MaxStop, the more specific constraint with scope over obstruent stops only, must be ordered before NasHar, which applies to all segments. The PT-dictated interaction guarantees the results illustrated in (13) for Malay and similar languages.

(13) Obstruent stop opacity in Malay

| | Input: makan | MaxStop | NasHar |
|---|---|---|---|
| ☞ | a. mãkan | | *** |
| | b. mãkãn | *! | |

There is reasonable doubt as to whether (13b) is a licit candidate, because it contains a phonetically illicit segment. However, the question of its legitimacy is moot, because the violation of MaxStop renders it non-optimal.

The validity of any proposed constraint is enhanced, if it can be justified independently of a single phenomenon. In other words, while the satisfaction of MaxStop happens to account for the opacity of obstruent stops to nasal harmony, the analysis would be strengthened, if this constraint could be shown to play a role in the explanation of some other phenomenon. I argue that it does; MaxStop helps to define the content of consonant inventories. In accordance with current thinking, we may assume that all consonants types (i.e., stops, fricatives, nasals, liquids, etc.) are inputs in all grammars and constraints determine the set that constitutes the inventory of a particular language. MaxStop demands that stops in the input be preserved as the same segment type in the output. Since stops are present in every language (Ladefoged & Maddieson 1996; Lass 1984), MaxStop must be part of every grammar and, therefore, fits the profile of a principle rather than a parameter of variation. This explanation for the universality of stops also explains why there could not be a language that lacks stops. A ban on this class of segments would have to be attributed to the satisfaction of a hypothetical constraint no stop (*stop), and the theory rules out the possibility of a constraint that would be completely antagonistic to MaxStop. The postulation of the latter constraint as a grammatical principle ensures that the ranking illustrated in (13) is imposed on every nasal harmony system.

The presence of MaxStop in every grammar and its dominance over NasHar also entail that nasal harmony cannot be facilitated by substituting a nasal for an oral stop. In other words, there could not be a case where NasHar commands the selection of [mãŋãn] (14b) as the realization of underlying *makan*.

(14)

| | Input: makan | MaxStop | NasHar |
|---|---|---|---|
| ☞ | a. mãkan | | *** |
| | b. mãŋãn | *![9] | |

In disfavoured (14b) the input obstruent differs from its surface correspondent in voicing, sonorancy and nasality. Any one of these differences is sufficient to render this candidate sub-optimal. The universal exclusion of the candidate with a nasal substitute for an underlying obstruent stop is significant, because nasal substitution is otherwise well documented in the phonological literature as a property of a diverse group of languages. Phonological theory must therefore allow for the possibility of nasal substitution but disallow such an option when it is induced merely to facilitate the extension of nasality.[9]

Another process that could never be exploited to facilitate nasal harmony is debuccalization. This phenomenon takes the form of the replacement of an underlying obstruent stop by a glottal stop; it occurs, for example, in Kelantan Malay (Onn 1976). Since nasal harmony, once activated, readily penetrates laryngeals, the substitution of a glottal stop for an obstruent stop has the potential to eliminate an impediment to nasal spreading. Nevertheless, there are no attested cases where an underlying string like *makan* is realized as [mãʔãn]. Obviously, the latter representation is not absolutely illformed; the comparable sequence [maʔap] 'pardon' occurs in Malay. Therefore, the sub-optimality of [maʔan] is relative to the input *makan*. The relation characterized in (15a) is sanctioned by all nasal harmony systems, but the one in (15b) is universally unattested.

(15) a. maʔan > mãʔãn
b. *makan > mãʔãn

In (15b) the underlying medial stop and its surface correspondent differ in at least two ways. For one, the glottal stop is a placeless segment, while its underlying counterpart is specified as velar. The two also differ in nasality. The obstruent stop is oral, but, when nasalization extends through a glottal stop, the resulting segment is both phonetically and phonologically nasal (cf. Walker & Pullum 1999). Consequently, candidate (16b) incurs at least two violations of MaxStop, and any violation of the latter to satisfy the requirements of lower-ranked NasHar is fatal.

(16)

| | Input: makan | MaxStop | NasHar |
|---|---|---|---|
| ☞ | a. mãkan | | *** |
| | b. mãʔãn | *! | |

The universal ranking MaxStop » NasHar rules out a third logical possibility for the realization of nasal harmony. It predicts that the deletion of an obstruent stop should not occur to facilitate the extension of nasality, and again the evidence confirms the correctness of the prediction. Although underlying stops are sometimes absent from surface representations (e.g., French, Maori), none of the documented cases of nasal harmony reveals a pattern in which the hypothetical input /makan/ is realized as [mããn].

(17)

| | Input: makan | MaxStop | NasHar |
|---|---|---|---|
| ☞ | a. mãkan | | *** |
| | b. mããn | *! | |

The loser (17b) incurs multiple violations of MaxStop, since all input features are missing from the output.

### 3.2 Explaining segment transparency in nasal harmony

Since stops cannot be penetrated by nasality to best satisfy the requirements of nasal harmony, a universal no-gap condition should prohibit the skipping of this class of segments. This prediction is generally borne out; as a class, obstruent stops are never transparent. Nevertheless, we know that voiceless stops may be skipped. Piggott & van der Hulst (1997) provide a solution to the transparency problem. A crucial component of their analysis is the hypothesis that the nasal feature may be licensed as a property of either a segment or a syllable. The theory of licensing that underlies this hypothesis is spelled out in Piggott (1997, 1999, 2000).[10]

Features like nasality are always attributes of segments. This link between segments and nasality is captured by the following licensing constraint.

(18) NASAL LICENSING/SEGMENT (NasLic/Seg)
Nasal is licensed in every occurrence as a property of a segment.

The satisfaction of this constraint requires that, when the feature nasal is present, it must be a segmental property. However, in several languages, individual instances of nasal cannot be limited to single segments but must minimally encompass a vowel and any tautosyllabic voiced segment. Nasal-oral contrast, in such cases, is evidently between syllables. Among the languages where this contrast is attested are Yoruba (Pulleyblank 1988), Isekiri (Omamor 1979), Jukun (Welmers 1973), Gbe (Capo 1981), Kaingang (Wiesemann 1972) and Chaoyang (Yip 1996). Some examples of nasal syllables in Kaingang are given in (19a), while (19b) illustrate the oral type. Note that full nasals occur only in nasal syllables, while their counterparts in oral syllables are voiced prenasalized stops.

(19) Nasal syllables in Kaingang

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| a. | r̃ã | 'sun' | *rã | b. | ra | 'toward' |
| | ỹã | 'tooth' | *yã | | ya | 'already' |
| | w̃ãn | 'taquara' | *wãn | | wa | 'carrying' |
| | nã | 'lie' | *$^{\text{n}}$dã | | $^{\text{n}}$da | 'arrow' |
| | ỹãra | 'spit' | *ỹãr̃a | | $^{\text{ŋ}}$goyo | 'water' |
| | kur̃ã | 'day' | *kurã | | kara | 'all' |

Significantly, while all voiced segments in a Kaingang nasal syllable must be nasal, voiceless obstruents may be tautosyllabic with either oral or nasal vowels, as the following data show.

(20)

| | | | | | |
|---|---|---|---|---|---|
| a. | kutæ̃ | 'fall' | d. | kãtə | 'guts' |
| b. | hapæ̃ | 'good' | e. | ỹãtõ | 'dull' |
| c. | kɨšã | 'moon' | f. | sɨ̃ | 'small' |

To account for the realization of nasality in a language like Kaingang, we must first determine how the grammar provides for the nasal-oral contrast. I propose that it is a consequence of the satisfaction of the following constraint, the activation of which is language-particular.

(21) NASAL LICENSING/SYLLABLE (NasLic/σ)
Nasal is licensed in every occurrence as a property of a syllable.

Since features are always associated with segments at the phonetic level, phonological theory must provide a nasal syllable with an empirically distinctive profile. This is accomplished by the following requirement.

(22) Syllable Nasalization (SylNas)
Within a nasal syllable, all voiced segments must bear nasality.[11]

Voiceless stops and fricatives in nasal syllables are not forced to bear nasality and (generally) do not do so.

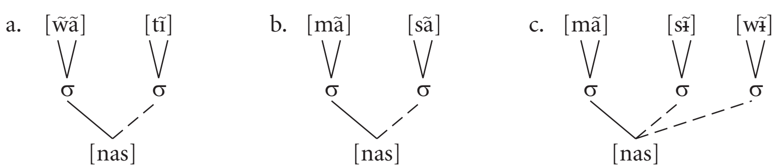
SylNas and NasLic/σ are structural constraints. As indicated earlier, the P&P framework incorporates a meta-condition that requires such constraints to outrank all members of the faith family. I now identify this meta-condition as *Structural Integrity*.

(23) Structural Integrity
structure » faith

In languages where nasal harmony is partly regulated by activation of (21) and (22), the following ranking of constraints must be in effect.

(24) NasLic/σ, SylNas » MaxStop » NasHar

This ranking commands violations of MaxStop but only as is necessary to meet the demands of SylNas. We would expect voiced stops and other voiced segments to be nasalized, but voiceless stops and voiceless fricatives would be unaffected. The neutrality of voiceless segments would not affect the extension of nasality, however, because this feature is transmitted from syllable to syllable. An informal illustration of the extension of nasality in this language with the ranking in (24) is provided below. The language is Southern Barasano, and the syllables are circumscribed by square brackets.

(25)



The two tableaux in (26) below show how the ranking in (24) selects optimal outputs for a pair of hypothetical inputs /mada/ and /mata/.

(26)

| | a. Input: mada | NasLic/σ | SylNas | MaxStop | NasHar |
|---|---|---|---|---|---|
| | i. [mã] [dã] | *! | *! | | |
| ☞ | ii. [mã] [nã] | | | * | |

| | b. Input: mata | NasLic/σ | SylNas | MaxStop | NasHar |
|---|---|---|---|---|---|
| ☞ | i. [mã] [tã] | | | | |
| | ii. [mã] [nã] | | | *! | |

The above demonstration assumes that, if NasLic/σ is satisfied, the demands of SylNas must also be met and vice-versa. The first candidate in (26a) is therefore disqualified, because the second syllable contains a voiced oral consonant. In comparison, the first candidate in (26b) satisfies the requirements of both NasLic/σ and SylNas. It will therefore always be superior to a candidate like (26bii) that incurs a violation of MaxStop.

The analysis in this section fully explains why segment neutrality in nasal harmony sometimes takes the form of transparency. We also understand why voiceless stops and voiceless fricatives exhibit transparent behaviour to the exclusion of all other segment types. Sonorants and voiced obstruents could never be skipped.

### 3.3 The violability of a principle

According to conventional thinking, principles are equivalent to constraints that cannot be violated by any language. The P&P framework defended here and elsewhere does not subscribe to such a tenet. A principle is considered to be a constraint that is satisfied in every grammar, if the relevant conditions are met. For example, SylNas is a principle that comes into play only when the activation of NasLic/σ provides for the occurrence of nasal syllables. Under the latter conditions, the requirements of SylNas must be satisfied. However, there is no theoretical imperative that a principle be completely inviolable. Notice that MaxStop qualifies as a principle. Nevertheless, we see that it is violated in languages with Type B nasal harmony to satisfy the demands of SylNas. The obvious implication is that violations of a principle may be commanded in order to satisfy the demands of a higher ranking constraint.

Violations of MaxStop are not necessarily limited to patterns in which nasals replace voiced obstruent stops. There are attested cases where nasals are substituted for voiceless stops. An example of such a case is provided by Chukchi, one of the languages of the Chukotko-Kamchatkan group.

(27) Chukchi nasal substitution (Odden 1994)

| | | | | |
|---|---|---|---|---|
| a. | rəpən | ‘flesh side of hide’ | rəmn-ət | ‘flesh sides of hides’ |
| b. | pəŋəl | ‘news’ | ɣa-mŋət-len | ‘having news’ |
| c. | təm-ək | ‘to kill’ | ɣa-nmə-len | ‘he killed’ |
| d. | rətən | ‘tooth’ | rənn-ət | ‘teeth’ |

e. təŋe-k 'to grow' ɣe-nŋe-lin 'it grew'

Each of the forms in the second column contains a nasal correspondent of an underlying voiceless obstruent stop. For example, the medial /p/ of the root /rəpən/ is replaced by the corresponding bilabial nasal [m] (27a).

In Chukchi, the nasal substitute is always pre-consonantal. Hence, the prima facie evidence would suggest that the context for the substitution is a coda position. Chukchi is therefore one of a number of languages (e.g. Ponapean, Korean, Japanese) that ban obstruent codas. For expository purposes, I refer to the constraint that enforces such a ban as NO CODA OBSTRUENT.

(28) NO CODA OBSTRUENT (*CODA-OBS)
An obstruent in a coda position is unlicensed.

This constraint expresses a restriction on syllable structure and must therefore be superordinate to MAXSTOP. (The illustration in the following tableau abstracts away from the role of syncope.)

(29)

| | Input: rəpən-ət | *CODA-OBS | MAXSTOP |
|---|---|---|---|
| ☞ | a. rəm.nə-t | | * |
| | b. rəp.nə-t | *! | |

The satisfaction of *CODA-OBS inevitably incurs a violation of MAXSTOP. Although the final consonants of the candidates in (29) play no role in the selection of the winner, we can be fairly confident that they are not coda segments and therefore incur no violations of *CODA-OBS. In comparison, the obstruent at the end of the first syllable of (29b) is fatal to this candidate.

The satisfaction of *CODA-OBS is sometimes achieved by substituting a glottal stop for an underlying obstruent stop. This pattern occurs in the Austronesian language, Agutaynen (Quakenbush 1991). When syncope in this language creates an opportunity for an obstruent stop to appear before a consonant, the resulting cluster surfaces with an initial glottal stop.

(30)

| | Root | Affixed form | |
|---|---|---|---|
| a. | petek | ma-petek [maʔtek] | 'will crack' |
| b. | petek | petek-en [peʔken] | 'crack!' |
| c. | teked | teked-en [teʔden] | 'tie!' |

The coda glottal stop meets the *CODA-OBS requirement but incurs a violation of the lower-ranked MAXSTOP.

Violations of MAXSTOP also occur in a common Austronesian pattern where stem-initial obstruents are sometimes replaced by nasals. These substitution patterns are usually morphologically governed (see footnote 10). A typical case involves the fusion of an underlying nasal-obstruent sequence into a simple nasal (e.g. N+p/b > m, N+t/d > n, N+k/g > ŋ). The following examples from Bonggi (Boutin 2000) illustrate this phenomenon.

(31) a. ŋ-bagi [mãgi] 'ACT.IND.SOA-divide something'[12]
b. ŋ-tɔluʔ [nõluʔ] 'ACT.IND.SOA-pursue'
c. ŋ-kɔrin [ŋõri$^{k}$ŋ] 'ACT.IND.SOA-dry something'

The nasal that results from fusion contains the place specification of the obstruent and the nasality of the nasal. Thus, it qualifies as the surface correspondent of both input segments. Each of the initial nasals in the surface forms of (31) is therefore a substitute for an underlying obstruent stop. The constraint that commands this substitution, referred to henceforth as NC-FUSION, applies to stem-initial stops, a subset of the segments to which MAXSTOP applies. Hence, its requirements take precedence over the demands of MAXSTOP.

Nasal harmony is also a feature of Bonggi phonology. The pattern is the Malay type; nasality extends through vowels and semivowels (32a) but is blocked by liquids and obstruents (32b).

(32) a. mien [mĩ̃ɛ̃n] 'aunt'
njawa [ɲãw̃ə̃] 'life, soul'
b. ŋ-tuluŋ [nuluŋ] 'ACT.IND.SOA-assist someone'
nagaʔ [nãgəʔ] 'dragon'

In Bonggi, the meta-conditions on ranking dictate the order in (33).

(33) NC-FUSION » MAXSTOP » NASHAR

This UG-imposed ranking will always favour a candidate in which a nasal harmony span encompasses the nasal correspondent of an obstruent stop, provided that the nasal is required to satisfy the dominant NC-FUSION constraint.

(34)

| | Input: ŋ-bagi | NC-FUSION | MAXSTOP | NASHAR |
|---|---|---|---|---|
| | a. mbagi | *! | | **** |
| | b. mãŋĩ | | **! | |
| ☞ | c. mãgi | | * | ** |

Candidate (34a) incurs a fatal violation of NC-FUSION. There are two violations of MAXSTOP in (34b) only one of which is induced by NC-FUSION. Hence this candidate must lose to (34c), showing again that, while MAXSTOP can be violated, nasal harmony itself cannot induce such a violation.

Additional evidence for the violability of MaxStop comes from the interaction of nasal harmony and lenition in Bonggi. According to Boutin (2000), underlying /k/ weakens to [h] at the beginning of an unstressed syllable in this language. Cross-linguistically, this is the favoured context for this type of consonant weakening.

(35) Lenition in Bonggi
  a. pakakas [ɸəkáhəs] 'tools, equipment'
  b. tikukur [tikúhʊr] 'spotted dove'

Since Bonggi lenition affects just one of the stops in only one of the contexts in which stops may be found, the constraint that commands the phenomenon, informally identified as lenition, is required by *Pāṇini's Theorem* to be superordinate to MaxStop (i.e. lenition » MaxStop). (A trochaic stress foot is identified by parentheses in the competing candidates.)

(36)

| | Input: tikukur | lenition | MaxStop |
|---|---|---|---|
| | a. ti(kúkʊr) | *! | |
| ☞ | b. ti(kúhʊr) | | * |

More significantly, the laryngeal correspondent of an underlying /k/ is opaque to nasal harmony in Bonggi.

(37) Laryngeal opacity in Bonggi
  a. manakan [mə̃nãhəᵈn] 'nephew, niece', *mə̃nãh̃ə̃n
  b. ŋ-akun [ŋãhʊᵈn] 'ACT.IND.SOA-admit something', *ŋãh̃ʊ̃n

The following tableau illustrates how the universal ranking of the relevant constraints selects the surface representation that corresponds to underlying /ŋakun/.

(38) Bonggi lenition and nasal harmony

| | Input: ŋakun | lenition | MaxStop | NasHar |
|---|---|---|---|---|
| | a. ŋãkuᵈn | *! | | *** |
| | b. ŋãh̃ʊ̃n | | **! | |
| ☞ | c. ŋãhʊᵈn | | * | *** |

The failure to satisfy the requirements of lenition dooms (38a). While lenition occurs in both (38b) and (38c), there is no tolerance for any violation of MaxStop that is not commanded by the demands of lenition. Therefore, the nasalization of [h] in (38b) is a fatal weakness. The ranking of the active constraints of Bonggi must converge on the selection of (38c).[13]

To summarize this sub-section, I have demonstrated that the principle which prohibits changes to underlying stops can be violated, but violations are only tolerated when they are induced by the satisfaction of constraint that is either more specific than MaxStop or qualifies as structural (e.g. *coda-obs). Moreover, nasal harmony cannot induce a correspondent of a stop to incur a violation of MaxStop, even when the correspondent is a segment that is otherwise susceptible to nasalization.

## 3.4 The opacity of fricatives

The fricative class is opaque to nasal harmony in most Type A systems. The postulation of a link between consonant opacity and consonant inventories leads us to attribute the opacity of fricatives to the satisfaction of the following constraint.

(39) Max Fricative (MaxFric)
An underlying fricative must have an identical surface correspondent.

MaxFric must outrank NasHar. The effect of the PT-dictated ranking is illustrated by the Sundanese example in (40).

(40)

| | Input: mĩʔãsih | MaxFric | NasHar |
|---|---|---|---|
| ☞ | a. mĩʔãsih | | *** |
| | b. mĩʔãs̃ĩh | *! | |

Candidate (40a) wins, because it better preserves the identity of the input fricative than does (40b) where the fricative is nasalized.

Like MaxStop, MaxFric also helps to determine the shape of consonant inventories. However, fricatives are not present in all languages; this class of consonants is absent from most Australian languages (Maddieson 1984; Hamilton 1996). MaxFric is therefore a parameter. When it is active, the consonant inventory of the language must include fricatives and these segments must display neutrality to nasal harmony, if present. The parameter status of MaxFric allows for the possibility that languages may opt to ban fricatives by invoking its antagonistic counterpart *Fric. However, languages that do not force fricatives to be represented in their inventories do not necessarily have to ban the occurrence of these segments. If they are not subject to a direct prohibition, fricatives might simply emerge from the set of input segments. In the latter scenario, an inactive MaxFric would have no effect on the occurrence of fricatives and could not require them to be neutral to nasal harmony. Applecross Gaelic exemplifies such a scenario; the emergent fricatives of this language are targets of nasalization. With regard to the status of fricatives, there are therefore three types of languages. One type must have fricatives because of the enforcement of MaxFric; it includes Sundanese and the other languages where fricatives are opaque to nasalization. A second type includes a language where MaxFric is inactive but fricatives emerge because they are not explicitly banned by *Fric; Applecross Gaelic exemplifies this type. The third type is, of course, the Australian languages where *Fric is active and fricatives are completely banned.

## 4. Limiting freedom of choice

The preceding explanation of the behaviour of fricatives predicts that other non-obligatory segment classes will also vary between being neutral and being targeted by nasalization. Liquids and glides, which do not have to be present in every consonant system, display such variable behaviour cross-linguistically. Their neutrality is evidence of the activation of the following constraints.

(41) a. Max Liquid (MaxLiq)
An underlying liquid must have an identical surface correspondent.
b. Max Glide (MaxGlide)
An underlying glide must have an identical surface correspondent.

When these constraints are inactive, liquids and glides, respectively, may be targets of nasalization. Glides, for example, are opaque in Sundanese but are nasalized in Malay.

If the choice to activate (or not activate) a parameter were freely exercised by languages, we would expect that the neutrality of segments other than stops to be completely unpredictable. The evidence contradicts this assumption. We can readily deduce from the patterns summarized in (6) that the neutrality of an optional consonant may be dependent on the neutrality of another. For example, if glides are opaque to nasal harmony, the other optional consonants (i.e. fricatives and liquids) must be also. Conversely, if fricatives and/or liquids are targeted, then glides cannot be neutral. I propose to link the dependencies displayed by segment neutrality to the relative markedness of consonant types. A number of papers, notably Rice (1992), Rice and Avery (1993) and Causley (1999), argue that markedness can be correlated with the amount of information that is needed to identify a class of segments. A consensus from this research is that sonorants are more marked than obstruents. Within the obstruent class fricatives are more marked than stops, while approximants (i.e. liquids and glides) are more marked than nasals. Finally, within the approximant class glides are more marked than liquids. These markedness relations converge on the following markedness scale for consonants, where the most marked is at the left edge.

(42) Consonant Markedness Scale
Glides » liquids » nasals » Fricatives » Stops

The above scale resembles the familiar sonority scale, but unlike the latter it does not include vowels, nor does it make any provisions for laryngeals, which are widely classified as non-consonantal.

From the representational characterization of markedness, Causley (1999) readily derives a notion of optimal segment; the less information a segment contains the better representative of its type it is. Thus, for example, nasals are better sonorants than liquids, and stops are better obstruents than fricatives. Indeed, stops must be the optimal consonants. Given such conclusions, one would expect UG to employ a range of strategies that favour the emergence of the better exemplars of consonants. We have already introduced one of these strategies. The designation of MaxStop as a principle ensures that there could not be a general injunction against the appearance of stops and guarantees the presence of the best consonant in every language. I postulate a second UG strategy. The latter requires that, when a language activates a constraint that commands the presence of a certain consonant-type in its inventory, it must respect the markedness scale in (42) and ensure that less marked types are also present. Such an outcome is achieved by imposing the following fixed ranking on the members of the family of constraints that command the presence of consonants in inventories.

(43) Consonant Preference Hierarchy (CPH)
MaxStop » (MaxFric) » (MaxLiq) » (MaxGlide)[14]

We deduce from this scale that the activation of a lower-ranked constraint automatically guarantees the activation of any member of the group that is more highly ranked. For example, when a language demands that liquids be present, it must also activate MaxFric to ensure that fricatives are also present. The implication is that, although MaxFric is a parameter of variation, the freedom to activate or not activate is limited by whether or not some lower-ranked constraint is active.

The Consonant Preference Hierarchy and *Pāṇini's Theorem* combine to control the manifestation of segment neutrality in nasal harmony systems.

*Pāṇini's Theorem* guarantees that any CPH constraint must outrank NasHar and the hierarchy determines the possible grouping of neutral segments.

(44) MaxStop » (MaxFric) » (MaxLiq) » (MaxGlide) » NasHar

The groupings are predictable by varying the exercise of the choice to activate one of the optional constraints of CPH. One group consisting of glides, liquids, fricatives and stops would be found in systems where MaxGlide is active; Sundanese displays such a system. When a language chooses instead to activate MaxLiq, the Malay pattern with neutral liquids, fricatives and stops is predicted. The decision to activate MaxFric limits the neutral group to fricatives and stops (e.g. Kolokuma Ijo). Finally, if a language decides not to activate any of the optional constraints of CPH, stops alone will function as neutral segment (e.g. Applecross Gaelic).

Notice that this analysis of segment neutrality in nasal harmony makes a very strong and testable empirical prediction about the relationship between this phenomenon and the content of a consonant inventory. It predicts that whenever a particular segment-type is neutral the consonant inventory must contain representatives of all less marked types. In other words, a language where liquids are neutral to nasal harmony but fricatives are absent from the inventory should be impossible. The predicted correlation between segment neutrality and the obligatory consonants of inventories is summarized in (45).

(45)

| Neutral segments | Content of inventories |
|---|---|
| Glides | Stops, fricatives, liquids, Glides |
| Liquids | Stops, fricatives, liquids |
| Fricatives | Stops, fricatives |

If there was no relationship between inventories and segment neutrality, optional consonant-types should be randomly distributed across languages with nasal harmony. For example, since fricatives are cross-linguistically optional, we might expect to find a language that lacks this class of segments but manifests a nasal harmony pattern with neutral liquids and stops. No such language has been attested.

## 5. Theoretical implications

All of the observations about segment neutrality that are addressed in this paper have been recognized in other analyses of nasal harmony. To appreciate the superiority of the framework and the explanations that it generates, we must consider merits of competing alternatives. The appropriate alternative is an analysis in the currently dominant framework of Optimality Theory. Such an analysis is presented by Walker (1998). One of the differences between Walker's proposal and the analysis presented in this paper is in the identification of the constraints that command segment neutrality. Walker postulates a set of restrictions on the feature combinations instead of the requirements that the identity of underlying segments must be preserved. Although this difference between the two analyses is important, it is not crucial. The more significant differences are in the P&P and the standard OT frameworks. Both frameworks accept that constraints must be ranked. It is the language-particular nature of constraint ranking in the OT framework that undermines its ability to explain all restrictions on the manifestations segment neutrality, especially the behaviour of obstruent stops.

Language-particular ranking of NasHar and MaxStop (or its equivalent) must allow for the targeting of obstruent stops in order to satisfy the demands of nasal harmony. The OT framework makes the counterfactual claim that stops may either be replaced by nasals or glottal stop or be deleted to facilitate the occurrence of nasal harmony. The random ranking of constraints makes it impossible to capture the fact that the demands of nasal harmony must always be subordinated to the need to preserve the identity of underlying stops. Notice that the problem of non-permeability of stops cannot be solved by assigning MaxStop the dubious status of an undominated and inviolable constraint; it is readily violated. Only a framework in which constraint-ranking is intrinsic to the nature of the constraints themselves captures the relative non-violability of a constraint like MaxStop. This means that differences between grammars cannot be reduced to differences between the ranking of the same set of constraints. The alternative defended in this paper is that grammatical differences are best explained as parametric choices in the activation of constraints.

## Notes

* The research for this paper was supported by a grant from the Social Sciences and Humanities Research Council of Canada (410-2001-0668).

**1.** See Prince and Smolensky (1993: 1) for some of the most important references.

**2.** Scobbie (1991, 1993) also incorporates the Elsewhere Condition into his theory of Declarative Phonology.

3. Trivial cases that appear to falsify the PT prediction can undoubtedly be found in the OT literature. However, a serious commitment to a restrictive theory where intrinsic ordering is favoured over arbitrary extrinsic ordering forces a re-evaluation of all such cases.

4. So-called cases of free variation would be attributed to co-existing grammars.

5. The parenthesized nasal is assumed to be in the underlying representation.

6. Examples of non-nasal words may contain nasal contours. These contours vary with purely oral stops, especially in intervocalic positions.

7. This paper offers no explanation for the restrictions on nasalization when the source is within a cluster.

8. This characterization of nasal harmony obviously has to be supplemented to allow for variation in directionality. Harmony may be progressive, regressive or even bidirectional.

9. Since there is more than one difference in features between the underlying obstruent stop and its surface correspondent, the losing candidate actually incurs more than one violation of MaxStop. However, for expository reasons only one violation is indicated here, since one is sufficient to doom the candidate.

10. Some of the Autronesian cases of nasal substitution appear to contradict this claim. For example, Blust (1997) cites data from Ngaju Dayak showing the replacement of an obstruent stop by a homorganic nasal that appears to initiate nasal extension (e.g. *pukul* 'a blow': *mã-mũkul* 'to hit'). However, such cases are restricted to stem-initial consonants and are almost certainly linked to the representation of the prefix. Notice also that nasalization does not trigger nasal substitution for the stem-medial obstruent in *mã-mũkul*.

11. See also Harris (1994, 1997) and Steriade (1995).

12. The designation of voiced segments as bearers of nasality may be linked to the fact that the typical nasals are sonorants and sonorants are normally voiced. The latter observation is exploited by Itô, Mester and Padgett (1995) in their explanation of postnasal voicing. As far as I am aware, only one language, Epena Pedee (Harms 1985), displays a property resembling syllable nasalization and allows for nasalization of a voiceless fricative. To accommodate such a possibility, the theory would have to be amended to allow nasality to be borne by all potential nasal-bearers within a nasal syllable.

13. According to Boutin (2000), ACT means 'actor' and IND.SOA means 'induced state of affairs'.

14. This conclusion provides some insight into the observation by McGinn (1979) that glottal stop is opaque to nasal harmony in the Austronesian language, Rejang; the segment might be the correspondent of an underlying stop. However evidence presented is not conclusive.

15. The fully elaborated scale would also include MaxNasal.

## References

Barnes, J. (1996). Autosegments with three-way contrasts in Tuyuca. *International Journal of American Linguistics, 62*, 31–58.

Blust, R. (1997). Nasals and nasalization in Borneo. *Oceanic Linguistics, 36* (1), 149–179.

Boutin, M. (2000). Blocking nasal spread in Bonggi. Unpublished Ms., Institut Linguistik SIL, Malaysia.

Capo, H. C. (1981). Nasality in Gbé: a synchronic interpretation. *Studies in African Linguistics, 12*, 1–43.

Causley, T. (1999). Complexity and markedness in Optimality Theory. Doctoral dissertation, University of Toronto.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Cohn, A. (1987). A survey of the phonology of the feature [+/-nasal]. Unpublished Ms., UCLA.

Gregores, E. & Suarez, J. A. (1967). *A Description of Colloquial Guarani*. The Hague: Mouton.

Hamilton, P. (1996). Phonetic constraints and markedness in the phonotactics of Australian aboriginal languages. Doctoral dissertation, University of Toronto.

Harms, P. (1985). Epena Pedee (Saija) nasalization. In R. Brend (Ed.), *From Phonology to Discourse: Studies in Six Colombian Languages* (pp. 13–18). Dallas: Summer Institute of Linguistics.

Harris, J. (1994). *English Sound Structure*. Oxford: Blackwell.

Harris, J. (1997). Licensing inheritance: an integrated theory of neutralization. *Phonology, 14*, 315–370.

Itô, J., Mester, R. & Padgett, J. (1995). Licensing and underspecification in Optimality Theory. *Linguistic Inquiry, 26*, 571–613.

Jones, W. & Jones, P. (1991). Barasano syntax. University of Texas-Arlington: Summer Institute of Linguistics.

Kiparsky, P. (1973). Elsewhere in phonology. In S. Anderson and P. Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 93–106). New York: Holt, Rinehart and Winston.

Kisseberth, C. (1970). On the functional unity of phonological rules. *Linguistic Inquiry, 1*, 291–306.

Ladefoged, P. & Maddieson, I. (1996). *Sounds of the World's Languages*. Oxford: Blackwell.

Lass, R. (1984). *Phonology: An Introduction to Basic Concepts*. Cambridge: Cambridge University Press.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

McCarthy, J. & Prince. A. (1995). Faithfulness and Reduplicative Identity. In J. Beckman et al. (Eds.), *Papers in Optimality Theory* (pp. 249–384). University of Massachusetts-Amherst.

McGinn, R. (1979). Outline of Rejang Syntax. Doctoral dissertation, University of Hawaii.

Odden, D. (1994). Adjacency parameters in phonology. *Language, 70*, 289–330.

Omamor, A. P. (1979). A phonological sketch of Isekiri. *Afrika und Ubersee, 62*, 190–223.

Onn, F. (1976). Aspects of Malay Phonology and Morphology. Doctoral dissertation, University of Illinois.

Paradis, C. (1988). On constraints and repair strategies. *The Linguistic Review, 6*, 71–97.

Penner, E. (1994). L'Harmonie nasale en Guarani: une approche autosegmentale. Thèse de Doctorat, Ecole des Hautes Etudes en Sciences Sociales.

Piggott, G. (1992). Variability and feature dependency: The case of nasality. *Natural Language and Linguistic Theory, 10*, 33–77.

Piggott, G. (1997). Licensing and alignment: A conspiracy in harmony. *Phonology, 14*, 437–477.
Piggott, G. (1999). At the right edge of words. *The Linguistic Review, 16* (2), 143–185.
Piggott, G. (2000). Against featural alignment. *Journal of Linguistics, 36*, 82–129.
Piggott, G. L. & van der Hulst, H. (1997). Locality and the nature of nasal harmony. *Lingua, 103*, 85–112.
Prince, A. & Smolensky, P. (1993). Optimality Theory. Unpublished Ms., Rutgers University and the University of Colorado at Boulder.
Pulleyblank, D. (1988). Vocalic underspecification in Yoruba. *Linguistic Inquiry, 19*, 233–270.
Quakenbush, S. (1991). Agutaynen glottal stop. Paper presented to the 6th International Conference on Austronesian Linguistics. Honolulu, Hawaii, May 20–24, 1991.
Rice, K. (1992). On deriving sonority: a structural account of sonority relations. *Phonology, 9*, 61–99.
Rice, K. & Avery, P. (1993). Segmental complexity and the structure of inventories. *Toronto Working Papers in Linguistics, 12* (2), 131–153.
Robins, R. (1957). Vowel nasality in Sundanese. *Studies in Linguistic Analysis* (pp. 87–103). London: Blackwell.
Scobbie, I. (1991). Attribute value phonology. Ph.D. thesis, University of Edinburgh.
Scobbie, I. (1993). Constraint violation and conflict from the perspective of Declarative Phonology. In C. Paradis and D. La Charité (Eds.), Constraint-Based Theories in Multilinear Phonology. *Canadian Journal of Linguistics* (Special Issue) *38*, 155–169.
Smith, C. & Smith, R. (1971). Southern Barasano phonemics. *Linguistics, 75*, 80–85.
Steriade, D. (1995). Underspecification and markedness. In J. Goldsmith (Ed.), *Handbook of Phonological Theory* (pp. 114–174). Oxford: Basil Blackwell.
Ternes, E. (1973). *Phonemic Analysis of Scottish Gaelic*. (Forum Phoneticum I). Helmut Hamburg: Buske Verlag.
Walker, R. (1998). Nasalization, neutral segments and opacity effects. Doctoral dissertation, UC-Santa Cruz.
Walker, R. & Pullum, G. (1999). Possible and impossible segments. *Language, 75*, 764–780.
Welmers, W. (1973). *African Language Structures*. Berkeley: University of California Press.
Wiesemann, U. (1972). *Die Phonologische und Grammatische Struktur der Kaingang Sprache*. The Hague: Mouton.
Williamson, K. (1965). *A Grammar of the Kolokuma Dialect of Ijo*. Cambridge: Cambridge University Press.
Yip, M. (1996). Lexicon optimization in languages without alternations. Ms., University of California-Irvine.

# Towards a theory of fundamental phonological relations*

Charles Reiss
Concordia University, Montréal

## 1. Introduction

In this paper I attempt to develop the "adequate formal account of identity references" that Odden (1988) demonstrates is necessary for phonological theory. As Odden points out

> ...languages differ in what constitutes 'identical' segments. Biblical Hebrew identical consonant fusion requires reference to complete identity (including voicing). Syrian Arabic allows identity to ignore pharyngealization and voicing, Koya allows identity to ignore retroflexion, and Telugu Syncope requires only rough identity computed at the place of articulation, which ignores voicing and narrow place distinctions such as alveolar/retroflex/palatal [461].

We will look at some of this data below in Section 2, which discusses the status of the Obligatory Contour Principle (OCP) from both an empirical and methodological perspective.

Again following Odden, I assume that "It is misguided to attribute every accidentally true statement about human language [or particular human languages – cr] to UG, for doing so trivializes the theory of UG itself"(461). Thus, linguistic theory should attempt to unify diverse phenomena by analysing them at an appropriately abstract level, instead of merely cataloging observations. In this spirit, I follow up on Odden's groundbreaking work, and related observations by Archangeli & Pulleyblank (1994), to propose that phonological theory needs the power provided by the existential quantifier and the universal quantifier to express identity references as conditions within the structural description of rules. Section 3 develops the notion of identity and nonidentity conditions in rules, and in Section 4. I propose and justify a new notation for expressing such conditions. My development of Odden's work serves as a small step towards making explicit what formal apparatus phonology must have access to.

Building on these results, I show that the need for quantificational statements entails the rejection of feature geometry in phonological representation. Feature geometric representation is insufficiently powerful, and must be replaced by an algebraic form of representation that allows the use of variables and indices for the purposes of identity checking.

An immediate benefit of approaching phonology in this minimalist fashion is that it helps us to discover, in Section 5, that certain *a priori* plausible rule- (or constraint-) types are actually unattested. In 5.2, I discuss this kind of 'overgeneration' and propose an explanation for such gaps in attestation.

## 2. The dubious status of the OCP

McCarthy (1986) discusses data from several languages in which a vowel which is expected for independent reasons to be deleted, is instead preserved if its deletion would cause identical consonants to be adjacent: Biblical Hebrew /ka:tab-u:/ → [ka:θvu:] but /sa:bab-u:/ → [sa:vavu:] because deletion would bring together the two underlying [b]'s (both of which are spirantized by an unrelated process).[1] The 'failure' of the deletion rule to apply is dubbed *antigemination* by McCarthy, since the rule is 'blocked' if its application would produce a geminate. McCarthy invokes the Obligatory Contour Principal (OCP) as the constraint which blocks the rule from applying. This phenomenon involves the failure of deletion rules just in cases where the rule would result in a string of identical adjacent consonants.

Yip (1988) provides a very useful summary, elaboration, and discussion of McCarthy's treatment of the OCP as a blocker of rules. Consider the following argument:

> If a language has a general phonological rule that is blocked just when the output would contain a sequence of identical feature matrices, we can conclude that the OCP is operating to constrain derivations ... The alternative is an ad hoc condition on such rules, as in [1]:
>
> (1) $A \rightarrow \emptyset / B\_\_C$
> Condition: $B \neq C$

> Such a condition not only incurs an additional cost (whereas the OCP is taken to be universal) but also lacks explanatory power, particularly if contexts *B* and *C* are necessary only to state the ad hoc condition.

In other words, Yip argues that a theory with language specific rules and a universal OCP is a better theory than one with language specific rules that correctly encode where the rule applies, because adding the necessary conditions to the statement of such rules makes them more complex.

Note that the examples that Yip mentions conform to the first (a) of the following three types of conditions on rule application, but Odden (1988) points out that in fact vowel syncope rules are found with all three of the following types of conditioning:

(2) Some conditions on vowel deletion rules (Odden 1988: 462)

a. Delete a vowel unless flanking Cs are identical.
b. Delete a vowel blindly [whatever the flanking Cs are].
c. Delete a vowel only if flanking Cs are identical.

Condition (a) can be restated as 'Delete a vowel if flanking Cs are *not* identical'. This is the condition described but rejected by Yip in (1) above: $B \neq C$. But note that Odden's type (c) condition would be written as follows:

(3) Odden's Condition (c) in the notation Yip rejects: $B = C$

In other words (a) demands nonidentity and (c) demands identity of segments in the structural description of a rule. Thus, there is no reason to propose, as McCarthy and Yip do, that rules that conform to condition (a) illustrate a universal principle of markedness – condition (c) is also a possible rule condition. A rule like (2c) *only* applies when it creates OCP violations – Odden refers to this phenomenon as *antiantigemination*. So a theory of *UG* must allow for both types. There is thus no good reason to claim that a universal principle, the OCP, *blocks* deletion in the (a) cases, since deletion can also be *required* in cases that lead to apparent OCP violations when a rule with conditions (b) or (c) applies. Stated in McCarthy's terms (although he does not mention such cases), deletion can be blocked (in case (c)) if the rule will *not* generate an OCP violation. This point was clearly made by Odden, though it seems to have been ignored in most of the subsequent literature.[2]

Note that the logic of attributing cases that fit the profile of (a) to a universal principle and ignoring cases that fit (c) is incoherent. Suppose we examine some data concerning a certain phenomenon and find that all cases fall into two categories, *x* or *y*. If we present only cases of *x* and proclaim that we have found that *x* is always true, then our claim is not valid, *no matter how many positive examples of x we adduce*. The existence of (c) cases, makes the existence of (a) cases uninteresting on their own. Odden's observations taken together *are* interesting, as we will see below. Simply put, case (c) is a counterexample to the claim that (a) is universal.[3]

## 2.1 Treating phonological pathology: The OCP as a rule trigger

The main point of Yip's paper is that the OCP not only *blocks* rule application as in McCarthy's antigemination cases, but also *triggers* it – it may be the case that a rule applies only to an input that violates the OCP. Instead of an argument based on formal simplicity in rule statements, as discussed above, Yip's discussion of the OCP as a rule trigger illustrates particularly well the assumption that the phonology repairs structures that are somehow pathological – ill-formed or marked or disfavored: "The main contribution of the OCP is that it allows us to separate out condition and cure. The OCP is a trigger, a pressure for change" (74).

In Yip's model the 'cure' is effected by language specific rules. In OT models that make use of similar constraints, the 'cure' emerges from the constraint ranking. Because of the violability of OT constraints, the winning candidate in an OT derivation is typically not fully 'cured' – certain marked structures may be present in the output form.[4] One goal of this paper is to work towards removing the notion of ill-formedness from the generative component of the phonology. There are representations that are generated, or formed, by grammars; there are representations that are not generated – that is, not formed; but there is no reason to believe that anything a grammar actually generates is ill-formed.

Yip provides a range of examples that show how different solutions can be applied to OCP violations. They include deletion, dissimilation, and assimilation rules (where assimilation represents multiple linking of a single node, and not identical adjacent nodes). One example of repair by deletion comes from Seri (Marlett & Stemberger 1983). This language has a rule that deletes a coda glottal stop in a syllable with a glottal stop in the onset:

(4) *Seri Glottal Stops*
   a. ʔa-a:ʔ-sanx → ʔ-a:-sanx 'who was carried'
   b. ʔi-ʔ-a:ʔ-kašni → ʔi-ʔ-a:-kašni 'my being bitten'
   c. koʔpanšx 'run like him!'

The rule only applies to tautosyllabic glottal stops so the second glottal stop in (4b) is not affected. In general, coda glottal stops can surface, as shown by (4c).

Yip's account of this process is the following:

> [We can] assume that the Laryngeal node is absent except for /ʔ/, and the entries for glottalization in [4a, b] are thus adjacent and identical and violate the OCP. This violation triggers a rule that operates in the domain of the syllable, and the language chooses [one of the possibilities for repairing OCP violations,] deletion of one matrix (either [+constricted] or [Laryngeal]). The actual rule has four parts, as shown in (5):
>
> (5) *Glottal Degemination*
> Domain: Syllable
> Tier: Laryngeal
> Trigger:
> Change: Delete second
>
> The environment is not stated, so the rule is unable to operate unless triggered "from the outside". The outside trigger is, of course, the OCP, a universal principle and thus free of charge.

In another example, Yip proposes that English uses epenthesis to 'cure' OCP violations of adjacent coronal stridents, thus accounting, for example, for the form of the plural morpheme after coronal stridents: *judges, couches, bushes, cases, etc.* In other words, if epenthesis did not apply, the adjacent coronal stridents would constitute an OCP violation. As Odden (1988) points out, the OCP is invoked rather opportunistically – note that it appears to be irrelevant to identity of adjacent [+voiced] specifications in words like *bins, rugs, hills, cars.* More seriously, Odden points out that there are rules that insert vowels only when doing so will specifically *not* repair an OCP violation. This is case (d) below. There are also rules that insert vowels regardless of the nature of the flanking consonants – case (e). And of course, there are rules that, like English epenthesis, depend on the total or partial identity of flanking segments – case (f).

(6) More conditions on vowel insertion rules (Odden 1988: 462)
 d. Insert a vowel unless flanking Cs are identical.
 e. Insert a vowel blindly [whatever the flanking Cs are].
 f. Insert a vowel only if flanking Cs are identical.

Parallel to (a), condition (d) can be restated as 'Insert a vowel if flanking Cs are *not* identical.' Thus there is no reason to see (f) as reflecting the OCP as a trigger when (d) shows that rules may be triggered if and only if they *fail* to fix OCP violations. The existence of rules with conditions (c) and (d) makes it unlikely that appealing to the OCP as either a trigger or blocker of rules is a fruitful endeavor.

## 3. The Identity and Nonidentity Conditions

More of Odden's data will be presented below. For now, note that it is equally possible for a rule to generate OCP violations (c) as it is to repair them (f). And it is equally possible for a rule to be 'blocked' from generating OCP violations (a) as to be blocked from fixing them (d).[5] Since the goal of phonological theory should be to define the set of computationally possible human languages, Odden's observations provide an excellent opportunity to study the purely formal nature of linguistic rules. In the following discussion, we will concentrate on syncope rules as a matter of expository convenience. Again, for expository convenience, we will refer to a schematic representation $C_1VC_2$. Odden's conditions (a) and (c) can be restated the following:

(7) The Nonidentity Condition on syncope rules (Version 1)
Delete a vowel if flanking Cs are *not* identical ($C_1 \neq C_2$).

(8) The Identity Condition on syncope rules (Version 1)
Delete a vowel if flanking Cs are identical ($C_1 = C_2$).

The apparatus of phonological representation must be at least powerful enough to express the Nonidentity Condition and the Identity Condition. This issue has implications for Feature Geometry as a model of phonological representation.

There is an insightful discussion of the need for Identity Conditions in Archangeli and Pulleyblank (1994: 368–373). These authors point out that "linked structures themselves are simply one type of configuration involving identity" (369). Archangeli & Pulleyblank present the 'Identity Predicate', a relation holding between two arguments, which "is important in a wide variety of phonological contexts" (369). In addition to the OCP cases, they cite the case of Tiv where [+round] spreads between vowels, if and only if they agree in height. Arguments against a linked structure analysis of identity conditions include cases where identity holds across a morpheme boundary – since the identical features belong to different lexical items, they cannot be stored as linked.

In the next section, I will formalize the identity condition and offer further arguments for the inadequacy of a 'linked structure' analysis of these conditions. Archangeli & Pulleyblank mention identity conditions holding of whole segments, as well as of individual features. We will see that it is also necessary to allow identity conditions over arbitrary subsets of the feature set. I will also show that linking is inadequate for the expression of non-identity conditions.

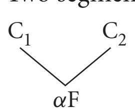
## 4. Feature algebra and conditions on rules

This section demonstrates that the notation of standard Autosegmental Representation (AR) is insufficiently powerful to represent the NONIDENTITY CONDITION and the IDENTITY CONDITION. I propose a solution to this problem by incorporating basic quantificational logic into an algebraic system of phonological representation called feature algebra (FA). I will argue that this system of representation has all the expressive power of Feature Geometry, as well as additional power that allows us to state hitherto unformulated aspects of phonological UG. The concern of this section will be with stating the structural descriptions (SDs) of rules. In other words, we will concentrate on describing the representations to which a rule applies, and not on the statement of the structural changes effected by the rule.

### 4.1 Shared feature values

One of the advantages of autosegmental theory, including Feature Geometry, is that it provides us with a visual representation of a situation in which two segments share a single specification for a given feature $F_n$. Such a situation can be a condition on the application of a rule (or the relevance of a constraint):

(9) Two segments linked to the same valued feature




Suppose that instead of representing this situation geometrically we did it algebraically with indices. For example, let $C_1$ and $C_2$ be understood as abbreviations for feature matrices such as the following:

(10) Segments as feature matrices

$$C_1 = \begin{bmatrix} (\alpha F_1)_1 \\ (\beta F_2)_1 \\ (\gamma F_3)_1 \\ \vdots \end{bmatrix} \qquad C_2 = \begin{bmatrix} (\delta F_1)_2 \\ (\varepsilon F_2)_2 \\ (\zeta F_3)_2 \\ \vdots \end{bmatrix}$$

$F_i$ denotes a feature, such as [nasal] and Greek letter variables denote the value ($\pm$) that feature $F_i$ has for a given segment.[6] The subscript outside of a pair of parentheses containing $\alpha F_i$ denotes the segment in question; thus, these subscripts are always 1 for $C_1$ and 2 for $C_2$.

If we want to express a state of affairs in which $C_1$ has the same value for some feature $F_n$ as another segment $C_2$, we can express this as follows:

(11) Identical values for $F_n$ using FA: $[(\alpha F_n)_1] = [(\beta F_n)_2]$

We thus express the fact that $C_1$ and $C_2$ have the same value for the feature mentioned. Perhaps we lose the visual metaphor of shared nodes, but the required *identity* condition on values is expressed by the equation. Obviously, this system can be extended to an arbitrary subset of the total set of features, even to the set of all features. We will do so below to formalize the Identity Condition, corresponding to Odden's condition (c).[7]

## 4.2 Indifferent feature values

An Autosegmental Representation may show two segments which are not linked with respect to a given feature. In such a case the standard interpretation is that such linking or lack thereof, is irrelevant to the application of the rule in question. The two segments *may* have identical values for a given feature, but this issue does not bear on the rule's applicability. This corresponds to Odden's condition (b). An example is a rule of schwa syncope in Hindi (Bhatia & Kenstowicz 1972). The form *daanəw + i* surfaces as *daanwi*, and *kaanən + i* surfaces as *kaanni*, showing that syncope is indifferent to identity or nonidentity of flanking consonants. The absence of association lines in the AR model is equivalent to the absence of an explicit algebraic statement of a relationship in our algebraic model. Since we constantly write rules which apply to *classes* of sounds in *classes* of environments, it is obvious that some feature values are irrelevant to the application of certain rules. For example, a rule that voices all stops between vowels does not refer to the various place of articulation features of the potential rule targets.

## 4.3 Obligatorily different feature values

Consider a rule which is only applied if two segments *disagree* with respect to some features. That is, consider rules conforming to Odden's conditions (a) and (d). For now let's consider a simple case where the two segments must disagree with respect to a single feature $F_n$. This can be represented trivially by overtly specifying the two segments, one as $\alpha F_n$ and the other as $-\alpha F_n$:

(12) Autosegmental representation of two segments with distinct values for a feature F

$$\begin{array}{cc} C_1 & C_2 \\ | & | \\ \alpha F & -\alpha F \end{array}$$

In algebraic terms this is easy to represent:

(13) Distinct values for $F_n$ using FA: $[(\alpha F_n)_1] \neq [(\beta F_n)_2]$

The algebraic formulation (13) expresses the fact that $C_1$ and $C_2$ have the opposite value for the feature mentioned. Again, it is obvious that this system can be extended to an arbitrary subset of the total set of features. (It can be extended to the whole set, but appears not to be – see below.)

## 4.4 The extra power of feature algebra with quantifiers

Note, however that there are examples of rule conditions (a) and (d), those that require that two segments be distinct, that cannot be expressed using just feature geometric association lines or feature algebra as sketched thus far. For example, imagine a requirement that $C_1$ and $C_2$ be different with respect to some arbitrary feature, that is any feature, or any feature out of a predefined subset of all the features. In other words, the two segments must *not be identical*, but it doesn't matter how they differ.

Let **F** be the set of all features. In order to express such a NONIDENTITY CONDITION we can make use of the existential quantifier:

(14) The NONIDENTITY CONDITION in FA (defined over all features)
$\exists F_i \in \mathbf{F}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$
There is at least one feature for which $\text{segment}_1$ and $\text{segment}_2$ have different values.

That is, there is some feature for which the two segments have a different value. Note that there is no way to represent NONIDENTITY using just Autosegmental Representation. This is because nonidentity can be due to a disagreement with respect to *any* arbitrary feature. The essence of autosegmental notation is the way in which it provides a geometric model of phonological structure. Being geometric, autosegmental representation does not make use of variables. Therefore, autosegmental notation is not sufficiently powerful to express non-identity conditions.

The requirement of difference can also be restricted to a subset of the features, $\mathbf{G} \subseteq \mathbf{F}$, for example, to the place of articulation features. Then, the more general version of the NONIDENTITY condition is the following:

(15) The NONIDENTITY CONDITION in FA (Final version)
$\exists\, F_i \in \mathbf{G}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$
For some specified subset of the features, there is at least one feature for which segment$_1$ and segment$_2$ have different values.

As we will see below, such a condition is necessary for the formulation of some well known phonological processes.

Once we admit the necessity of quantificational statements in our phonology, we can see that conditions of identity can also be expressed in such a fashion. Total identity can be expressed as follows:

(16) The IDENTITY CONDITION in FA (defined over all features)
$\forall\, F_i \in \mathbf{F}\ [(\alpha F_i)_1] = [(\beta F_i)_2]$
For all features, segment$_1$ and segment$_2$ have the same value.

whereas partial identity can be expressed by defining a subset of features, $\mathbf{G} \subseteq \mathbf{F}$ over which identity must hold. Total identity is just a special case of partial identity, where $\mathbf{G} = \mathbf{F}$.

(17) The IDENTITY CONDITION in FA (Final version)
$\forall\, F_i \in \mathbf{G}\ [(\alpha F_i)_1] = [(\beta F_i)_2]$
For some specified subset of the features, segment$_1$ and segment$_2$ have the same values.

Again, we shall see that such conditions are part of phonology.

## 4.5 Examples of conditions on identity and nonidentity

Note that McCarthy's account of antigemination, which uses the OCP to block rule application, involves a 'lookahead' effect: the phonology must see what the outcome of the rule *would be* and then 'decide' whether or not the rule is to be applied. In effect, the rule must be done and undone if the outcome is not satisfactory. An alternative to the rules-and-constraints lookahead solution is to build into the rule the conditions on its application. Note that this condition is just a part of the rule's Structural Description (SD), and a SD is needed in any event. In the case of antigemination, if the OCP is dispensed with, there is no lookahead, but instead a Nonidentity Condition is built into the rule. McCarthy's rule deletes a vowel in the environment $\#CVC_1\_\_C_2V$, unless it is blocked by the OCP. I propose replacing McCarthy's rule with one that deletes the vowel in the environment $\#CVC_1\_\_C_2V$ if $\exists\ F_i \in \mathbf{F}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$.

### 4.5.1 *The* Nonidentity Condition

All the examples of OCP blocking cited by McCarthy, including the Biblical Hebrew case illustrated on page 222, can be restated as rules with a condition 'apply unless two segments are identical'. Again this is equivalent to 'apply only if two segments are different, that is, non identical'. So these rules all exemplify the Nonidentity Condition. Odden and McCarthy also provide an example of antigemination from Iraqi Arabic:

(18) Antigemination in Iraqi Arabic (from Odden 1988: 452)

a. xaabar 'he telephoned' xaabr-at 'she telephoned'
ħaajaj 'he argued' ħaajij-at 'she argued'

b. *Syncope*
$V \rightarrow \emptyset\ /\ V(C)C\_\_\_\_CV$

The syncope rule applies normally in the form *xaabrat*, but is blocked, according to McCarthy, in ħ*aajijat* to avoid generating an OCP violation.

Under the theory developed here, rule application is not blocked by the OCP, but rather, the rule's SD includes a nonidentity condition. Rule 19 shows the necessary indexing of the consonants in the structural description:

(19) Revised Iraqi rule
$V \rightarrow \emptyset\ /\ V(C)C_1\_\_\_\_C_2V$
if $\exists\ F_i \in \mathbf{F}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$

To reiterate, the syncope rule is not written in an overgeneral form and *blocked* by the OCP, but instead the rule contains a condition (part of the structural description) which determines that the rule only applies when the consonants on either side of the vowel differ with respect to at least one feature. Since structural desciptions are obviously needed to determine where rules apply, there is no reason to decide arbitrarily that, say, the presence of flanking consonants should be part of the structural description, but that their identity or non-identity should not be.

The Cushitic language Afar, dicussed by McCarthy (1986) and Yip (1988), based on data found in Bliese (1981) provides a similar case of antigemination which is sensitive to a nonidentity condition. Vowels which appear under stress in the left hand column of (20b) are deleted in the related forms to the right, since the stress has shifted. The unstressed vowels in (20c) do not delete. The rule is stated in (20a) with the relevant nonidentity condition, and the notation [-stress].

(20) *Syncope in Afar*

a. $\underset{[\text{-stress}]}{V} \rightarrow \emptyset$ / $\#CVC_1$____$C_2V$ if $\exists\, F_i \in \mathbf{F}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$

| | | | |
|---|---|---|---|
| b. | xamíla | xamlí | ‘swampgrass’ (acc/nom-gen) |
| | ʕagára | ʕagrí | ‘scabies’ (acc/nom-gen) |
| | digibté | digbé | ‘she/I married’ |
| | wagerné | wagré | ‘we/he reconciled’ |
| c. | miđáđu | miđađí | ‘fruit’ (acc/nom-gen) |
| | xararé | | ‘I, he burned’ |
| | danané | | ‘I, he hurt’ |

The alternations in the (b) forms show the deletion of unstressed vowels in open syllables. The first two lines show the relevance of stress – only unstressed vowels are deleted. The next two lines show that deletion does not occur when the syllable is closed. The (c) forms show that the rules do not apply when the flanking consonants are identical. In other words, the rule applies only between non-identical consonants.

As noted above, there is no way to refer to non-identity with respect to some arbitrary feature without making use of the existential quantifier. In the next subsection, I will show that the use of the universal quantifier is not only as good as a geometric representation to express identity conditions, but that in some cases, a geometric representation will be insufficient, so the quantificational formulation is the only one with sufficient power.

### 4.5.2 *The* Identity Condition

In this section, I repeat three of Odden's examples of deletion processes requiring identity between flanking segments. According to Odden, Sherwood (1986) motivates a rule in Maliseet-Passamaquoddy which deletes the short vowels /ə/ and /ă/ in doubly open syllables when flanking consonants are identical. Obviously, this identity condition must be stated so that it applies to the whole feature set (or at least to those relevant to consonants). The Hebrew syncope rule we began with deletes vowels unless the flanking consonants are identical. We have encoded the condition 'unless identical' as 'necessarily nonidentical', using the existential quantifier and negation of identity. In the Maliseet-Passamaquoddy rule, a vowel deletes only when the flanking consonants are identical. This requirement can be encoded using the universal quantifier and the identity relation.[8]

Other rules mentioned by Odden, such as one posited by Jensen (1977) in Yapese, demand only homorganicity between flanking consonants, and not identity of Laryngeal or manner features. The Yapese rule deletes a vowel flanked by homorganic consonants if the first consonant is postvocalic or word-initial:[9]

(21) Yapese: syncope between homorganic consonants

a. $V \rightarrow \emptyset \,/\{V, \#\}\, C_1\underline{\quad}\# \, C_2$
if $\forall F_i \in \{[\text{coronal}], [\text{labial}], [\text{dorsal}]\}\ [(\alpha F_i)_1] = [(\beta F_i)_2]$

| | Underlying | Surface | Gloss |
|---|---|---|---|
| b. | ba puw | bpuw | 'it's a bamboo' |
| | ni te:l | nte:l | 'take it' |
| | rada:n | rda:n | 'its width' |

The Yapese data shows that we can use the universal quantifier to express identity. We now demonstrate that we should use the FA system with quantification, since Feature Geometry is insufficiently powerful.

### 4.5.3 *Another failing of Feature Geometry*

Odden cites data from Koya (Taylor 1969: 38) in which word final vowels are deleted if flanking consonants are identical, except that retroflexion is not used in the computation of identity. In other words, retroflex consonants group with plain coronals for the purposes of computing identity. The data and a FA formulation of the rule are given in (22).

(22) Koya: syncope between identical consonants – ignoring retroflexion

a. $V \rightarrow \emptyset \,/C_1\underline{\quad}\# \, C_2$
if $\forall\ F_i \in \{[\text{coronal}], [\text{labial}], [\text{dorsal}], [\text{spread glottis}], [\text{sonorant}], [\text{nasal}], [\text{lateral}]\}\ [(\alpha F_i)_1] = [(\beta F_i)_2]$

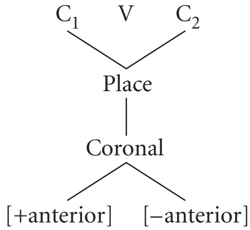
| | Underlying | Surface | Gloss |
|---|---|---|---|
| b. | na:ki ka:va:li | na:kka:va:li | 'to me it is necessary' |
| | a:ru ru:pa:yku | a:rru:pa:yku | '6 rupees' |
| | verka:ḍi digte | verka:ḍdigte | 'the cat got down' |

Rule (22a) is not particularly pretty, but it is correct (insofar as the set of features listed as relevant to consonant identity, retroflexion excepted, is correct) as a formalization of Odden's discussion.

It is extremely important to note that we have here further evidence for the inadequacy of Feature Geometry. As discussed above, Autosegmental Representation and Feature Geometry can be used to denote shared feature values, or even shared classes of features by a many-to-one mapping of segments (say, Root nodes) to feature structures. There are two implicit assumptions involved in using such representations. First, there is the assumption of locality: linked structures are adjacent in some sense, so that there are no intervening association lines to cross. In order to treat consonant identity across an intervening vowel as due to autosegmental linking, it is necessary to assume that the intervening vowel have no features which are shared with the two consonants. This assumption forces us to posit a different set of features for vowels and consonants, thus leading to the question of how the two types of segment can influence each other, or else it requires ad hoc segregation of vowels and consonants onto separate tiers. See Archangeli & Pulleyblank (1994: 368–70) for discussion.

Second, if nodes X and Y on a given tier both dominate a node A on a lower tier, then X and Y are assumed to be identically specified for node A (obviously) *and* any nodes that A dominates (by the transitivity of dominance). To illustrate, a structure like (23) is not well-formed since the [+anterior] and [-anterior] nodes are not ordered – temporal relations can only be determined with reference to the timing tier or Root tier.[10] In order to encode the ordering between these two specifications, one could mark them with an index to link them with $C_1$ and $C_2$ respectively. But in doing so, we have reverted to an algebraic notation.

(23) An illicit representation




But this view is incompatible with the situation in Koya where the retroflexion features are irrelevant to the computation of identity. Let's simplify matters and assume that the Koya vowels are not specified for Place features and thus flanking consonants can be considered adjacent. The syncope rule might then be assumed to apply when the flanking consonants share a Place node. However, the consonants should then be assumed to share all the dependents of the Place node. The Place node dominates the Coronal node and its dependents [anterior] and [distributed] in most models of Feature Geometry. However, the Koya rule applies even when the flanking consonants are not identically specified for these features. So, the rule cannot be stated using multiple linking of Place nodes or Coronal nodes.

If we take this as evidence that [anterior] and [distributed] are not dominated by Coronal, but are perhaps directly dominated by the Root node, we are just claiming that Feature Geometry contains less structure, that is we are arguing against its usefulness as a representation of the organization of features, and we are heading back to a model of unorganized feature matrices, as in (10). FA, in contrast to Feature Geometry, allows us to list all and only those features which are relevant to the Koya identity condition. We thus see that even in certain cases of the identity condition, FG representation is insufficiently powerful. It does not allow us to *exclude* from consideration of identity computations those nodes whose dominating nodes are *included* in the computation.

In order to show that this is not an isolated example, consider an additional case concerning place assimilation in modern Irish (pointed out to me by Morris Halle, p.c., discussed by Ni Chiosán and Padgett (1993) and Halle, Vaux and Wolfe (2000)). In this language, "word-final coronal nasals assimilate the primary place of articulation of a following stop, but crucially do not assimilate the secondary articulation of the stop. Palatalized [$n^j$] assimilates the Dorsal articulation of a following [g], but not its contrastive [+back] specification. Nonpalatalized [n] assimilates the Dorsal articulation of following [$g^j$], but not its [-back] articulation. In other words, the secondary articulation feature [back] does not spread whenever its supposedly dominating node spreads. The authors cited offer different solutions to this problem, while still maintaining the basic FG model. I take the data to be consistent with a rejection of FG, as argued for above.

Odden also mentions several cases of insertion rules that rely on an identity condition. For example, in Lenakel (Lynch 1978) schwa is inserted between identical consonants. Yip's examples of rules that break up (partially) identical clusters of consonants, such as the English epenthesis between coronal sibilants, can all be restated in terms of rules constrained by identity conditions.

## 5. Unattested Conditions

The previous section was really no more than a slight elaboration on Odden's important work on defining the range of possible conditions on rule application. In this section, I suggest that the use of FA actually helps us discover that two types of rule condition are unattested in phonology.

Phonological theory needs at least the power of the Nonidentity Condition and the Identity Condition. Interestingly, it appears not to need the power of the two conditions which are made by switching = and ≠ in the conditions already established. That is, no phonological rule appears to require what can be called the complete nonidentity condition (two segments must have opposite feature values for all of a given subset of features)[11] or the variable partial identity condition (two segments must be identical with respect to one member of a given subset of features, but it doesn't matter which particular member it is).

(24) Unattested: Complete Nonidentity Condition
$\forall F_i \in \mathbf{G}\ [(\alpha F_i)_1] \neq [(\beta F_i)_2]$
For some specified subset of the features, $\text{segment}_1$ and $\text{segment}_2$ have different values.

(25) Unattested: Variable Partial Identity Condition
$\exists F_i \in \mathbf{G}$ such that $[(\alpha F_i)_1] = [(\beta F_i)_2]$
For some specified subset of the features, there is at least one feature for which $\text{segment}_1$ and $\text{segment}_2$ have the same value.

An example of the Complete Nonidentity Condition would require that two segments have opposite values for, say, all place features, or even for *all* features. I know of no such case. For example, 'delete a vowel in the environment #CV$C_1$__$C_2$V if $C_1$ is [-anterior, -labial, +dorsal] and $C_2$ is [+anterior, +labial, -dorsal], or $C_1$ is [+anterior, -labial, +dorsal] and $C_2$ is [-anterior, +labial, -dorsal], *etc*'.

An example of the Variable Partial Identity Condition would require that two segments have the same value for some feature in a given subset: 'delete a vowel in the environment #CV$C_1$__$C_2$V if $C_1$ and $C_2$ are both [αanterior], or [αlabial], or [αdorsal], *etc*'.

In other words, we can trivially construct a condition using the universal quantifier and a nonidentity relation, or the existential quantifier and an identity relation, but it seems that such conditions never are needed by the phonology. In Section 5.2, I make a proposal concerning how this generalization should be treated by the theory.

## 5.1 Ambiguous cases

Obviously, there are rules for which the conditions on application *could* be expressed as an example of case (24) or (25), such as a rule requiring nonidentity for a single specific feature value. For example, Yiddish (Perlmutter 1988; Sapir 1915) has a rule that deletes the vowel in the plural suffix *-en* as long as the stem final consonant is not a nasal.[12] In terms of McCarthy's view of the OCP as a rule blocker, the rule fails to apply if it will bring two nasals into contact. Data appears in (26).

(26) Yiddish vowel deletion

| | Sg | Pl |
|---|---|---|
| 'language' | šprax | špraxn |
| 'ear' | oyer | oyern |
| 'magazine' | žurnál | žurnáln |
| 'sea' | yam | yamen |

Let's assume that the only relevant feature in conditioning the rule is [nasal], and that, for example, [sonorant] and [voice] are not part of the rule's structural description. Then, this rule *could* be stated in terms of a condition like (27), which is a highly restricted example of the Complete Nonidentity Condition: the set of relevant features contains the single feature [nasal], and the rule has the added condition that the second consonant be [+nasal].

(27) Delete the vowel between an onset consonant $C_1$ and a nasal $C_2$
if $\forall F_i \in \{[\text{nasal}]\}[(\alpha F_i)_1] \neq [(\beta F_i)_2]$

However, this condition can be expressed as well as an example of the NONIDENTITY CONDITION, as follows:

(28) Delete the vowel between an onset consonant $C_1$ and a nasal $C_2$
if $\exists\, F_i \in \{[\text{nasal}]\}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$

So it is not *necessary* to combine the universal quantifier with a nonidentity requirement to express the correct condition. Thus the claim that the COMPLETE NONIDENTITY CONDITION and the VARIABLE PARTIAL IDENTITY CONDITION are unattested in phonological rules as necessary kinds of conditions remains valid.

Other potential cases of the COMPLETE IDENTITY CONDITION arise when segments have redundant values for certain features, values that can be predicted on the basis of other features. Suppose a language **L** has in its consonant inventory only voiceless obstruents and voiced sonorants. Then a rule in **L** which apparently demanded nonidentity of voicing between two consonants could have the condition shown in (29):

(29) $\exists\, F_i \in \{[\text{voice}]\}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$
that is, $[(\alpha\text{voice})_1] \neq [(\beta\text{voice})_2]$

Alternatively, the correct (though extensionally equivalent) form of the condition might be one that referred to the feature [sonorant], as in (30)

(30) $\exists\, F_i \in \{[\text{sonorant}]\}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$
that is, $[(\alpha\text{sonorant})_1] \neq [(\beta\text{sonorant})_2]$

Finally, the COMPLETE IDENTITY CONDITION in (31) is also extensionally equivalent to the two versions of the NONIDENTITY CONDITION just listed.

(31) $\forall\, F_i \in \{[\text{sonorant}], [\text{voice}]\}[(\alpha F_i)_1] \neq [(\beta F_i)_2]$
that is, $[(\gamma\text{sonorant})_1] \neq [(\delta\text{sonorant})_2]$ AND $[(\varepsilon\text{voice})_1] \neq [(\zeta\text{voice})_2]$

Again, since we do not need to use the formulation in (31), and since we have reasons to believe that we never need to use such conditions, we have a tool for choosing among extensionally equivalent grammars. The phonology of **L** potentially has the condition in (29) or the one in (30), but definitely not the one in (31). This is the kind of argument suggested by Chomsky (1986) in refuting the claim of Quine (1972) that it is futile to attempt to choose among extensionally equivalent grammars: "Because evidence from Japanese can evidently bear on the correctness of a theory of $S_0$, it can have indirect – but very powerful – bearing on the choice of the grammar that attempts to characterize the I-language attained by a speaker of English" (Chomsky 1986: 38).

## 5.2 Towards an explanation for gaps in attestation

Assuming that the Complete Nonidentity Condition and the Variable Partial Identity Condition are really absent from human languages, how are we to treat this fact? One way to do so is to build the fact into UG as an explicit constraint against quantificational statements conforming to certain formats. This strikes me as the wrong way to approach the issue, if it merely consists of restating the descriptive observation as a principle of grammar and not being open to explanations outside of the realm of grammar. In this particular case, it may be possible to derive the gap from the relationship between language change and phonetics. Note that this approach is in no way incompatible with a nativist perspective – the nativist position is just that some (not necessarily *all*) non-trivial properties of language are innate.

Following work on the nature of sound change (Ohala 1990; Hale, forthcoming) and theoretical work in cognitive science (Pylyshyn 1984), Hale and Reiss (2000ab) argue that it is to be expected that attested patterns in the phonological systems of the world's languages reflect only a subset of what is computationally possible for the human phonological capacity.[13] In other words, all attested patterns must be generatable by the UG-given phonological capacity, but not all generatable patterns will arise, due to the nature of sound change and language acquisition. This point of view may be helpful in explaining why the Complete Nonidentity Condition and the Variable Partial Identity Condition are unattested.

In general, phonological processes arise diachronically from the reanalysis of sublinguistic (gradient) phenomena as grammatical (categorical, feature-based) phenomena. Now note that there is at least a partial correspondence between phonetics and phonology – for example, features referring to place of articulation tend to correspond to the nature of formant transitions between vowels and consonants. Work on feature detectors and the like, though far from complete, reflects the belief that we can study the nature of the traduction processes between phonetics (gradient phenomena) and feature-based phonology.[14] An identity condition defined over a subset of phonological features, therefore, will tend to be to some extent related to a 'natural class' of phonetic properties.

Similarly, a nonidentity condition on, say, segments in the environment of a rule implies identity of the segments in those environments in which the rule does *not* apply. In other words, nonidentity entails the existence of identity in the complement set of environments. From the phonetic/diachronic perspective, then, these two conditions are the same. These two types of condition depend, at least at the point when they are phonologized by a learner, on clusters of phonetic properties. However, from a synchronic, phonological perspective, they are computationally distinct – one requires the equivalent of universal quantification and identity; the other, existential quantification and non-identity.

In contrast to these two cases, it is hard to imagine how either the Complete Nonidentity Condition or the Variable Partial Identity Condition could be derived from definable clusters of shared phonetic properties. For example, there is no phonetic unity to be found between segment transitions involving [± voiced] agreement and those involving [±coronal] agreement. But this is the kind of phonetic phenomena required to give rise to the Variable Partial Identity Condition which requires agreement for an arbitrary feature among segments in a structural description.

It is also hard to imagine how having opposite values for a given set of features, as required by the Complete Nonidentity Condition, could lead to a phonetically stable pattern. Recall that the relevant case would allow *every* pairing (of members of the relevant set **G**) of opposite feature values. Restricting the relevant set to the members A and B, each line of (32) would instantiate an environment for rule application under a Complete Nonidentity Condition:

(32) Permutations of feature specification in the Complete Nonidentity Condition
[+A +B] *vs.* [–A –B]
[–A +B] *vs.* [+A –B]
[+A –B] *vs.* [–A +B]
[–A +B] *vs.* [+A –B]

It is hard to imagine how such sets of representations could correspond to a phonetically natural grouping.

To summarize, the Nonidentity Condition and the Identity Condition provide us with a lower limit on the computational resources of UG, whereas patterns of attestation reflect extragrammatical factors. I am not claiming that these conditions are in principle uncomputable by the phonological component of the mind, but rather that the nature of language transmission makes it unlikely, or perhaps even impossible, that they will arise.

## 6. Conclusions

I have argued that an algebraic formulation of phonological representation facilitates the incorporation of quantificational logic into structural descriptions. I have shown that well known cases of Nonidentity Conditions demonstrate the necessity of the existential quantifier (or its equivalent) for phonology. The cases of Identity Conditions that refer to arbitrary sets of features, sets that are not members of a class according to Feature Geometric models, support the claim that the universal quantifier, or its equivalent is also necessary. Since the Feature Algebraic notation is more powerful than Geometric notation, and since the Algebraic notation seems necessary, the Geometric notation can be dispensed with. This is not surprising since the original motivation for Feature Geometry now seems to have been somewhat misguided (see Hale & Reiss 2000ab). I have also argued that certain kinds of logically possible conditions, the Variable Partial Identity Condition and the Complete Non-Identity Condition appear not to be needed in the phonology.

If correct, the generalization about phonology using only conditions of type 15 and 17 may, as suggested, be derivable from principles outside the realm of grammar, and not reflect any real constraint on the nature of phonological computation. In any case, we can posit the Identity Condition and the Nonidentity Condition as *necessarily formulatable* by phonological *UG*.

Identity and non-identity are *symmetrical* relations, and thus appear to pose an immediate problem for the theme of this volume, the idea that all grammatical relations are asymmetrical. Like all claims, the necessity of these relations may be questioned in the future; however, I end with two suggestions for maintaining the insights derived from the asymmetry hypothesis. First, the hypothesis may need to be revised as follows. Note that the relation of c-command is non-symmetric: if α c-commands β, it may or may not be the case that β c-commands α. The relation of c-command has been replaced in syntax by *asymmetric* c-command. If we need such asymmetric relations, and we need symmetric relations like identity, then perhaps we can conclude that grammatical relations can be symmetric or asymmetric, but *not nonsymmetric*.

The second suggestion is that asymmetry is a property of the Universal Processor, not of Universal Grammar. Given a grammatical relation $R$ between α and β, the Processor evaluates $\alpha R\beta$, but not $\beta R\alpha$. Without bidirection evaluation, we could never have evidence for a nonsymmetric relation. This contrasts with symmetric and asymmetric relations for which the truth of $\beta R\alpha$ is predictable from the truth of $\alpha R\beta$. If α asymmetrically c-commands β, then β does not asymmetrically c-command α. If α is identical to β, then β is identical to α.

## Notes

* The author would like to thank audiences at the *2001 LSA Annual Meeting* in Chicago, the *Asymmetry Conference* at UQAM in 2001, the *Fourth Utrecht Biannual Phonology Workshop* in 2000, the MIT Phonology Circle, and the McGill Linguistics Department. Discussions with Brendan Gillon, Dana Isac, Mark Hale, Ida Toivonen and Ash Asudeh were particularly helpful.

1. It has been brought to my attention that vowel length in the Hebrew is actually difficult to determine. However, this issue is irrelevant to the point under discussion – any example of 'antigemination' will do and additional ones are provided below.

2. For example, Keer's (1999) recent OT thesis on the OCP, lists Odden's 1988 paper in the bibliography, but makes no reference to it in the text, even in sections discussing antigemination.

3. Providing a principled response to the reader who finds this discussion to constitute an argument for the violable constraints of Optimality Theory is beyond the scope of this paper, or perhaps even impossible, reducing to a question of faith.

4. We might refer to this idea as OT's Fallacy of Imperfection. Imperfection, or markedness, seems to be as irrelevant to linguistic theory as the notion of perfection.

5. Of course, (b) also potentially generates OCP violations, and (e) potentially repairs OCP violations.

6. I continue to refer to segments for expository convenience; however, the valued features belonging to a given segment are more accurately characterized as the valued features sharing an index. These indices, in turn are best understood as denoting association to elements of an X-slot or CV timing tier – valued features with identical indices are linked to identical elements of the timing tier.

7. A reviewer of an abstract of this paper complained that the formalization developed here merely restates Odden's original observation. I refer to Halle (1975: 532) for discussion of the importance of explicit, careful formalization: "[D]etailed concern for the formal machinery of phonology has led to significant insights into the relationship between superficially disparate facts ... [I]t has paid off in terms of a deeper grasp of the significance of certain empirical facts." My debt to Odden is, I assume, obvious.

8. We can remind the reader here that, as is always the case, only one of the two quantifiers is necessary, since they can each be derived from the other *via* negation. For example, '$\forall x, x = y$' is equivalent to '$\neg\exists x$ such that $x \neq y$'; and '$\exists x$ such that $x \neq y$' is equivalent to '$\neg\forall x, x = y$'. I continue to make use of both quantifiers for ease of exposition.

9. Presumably the correct generalization is that the first consonant is in an onset.

10. Such representations have been proposed for affricates – with a single Root node dominating both a [+continuant] and a [-continuant] specification. In such cases, one might propose that a default ordering principle ensures that the [-continuant] specification be ordered first, since that is how they are ordered in affricates. Such a strategy will not work in the present case, since the order of the features is not fixed in Koya.

**11.** That is 'complete' refers to all members of the given subset, not in general, to the whole feature set.

**12.** The rule is also sensitive to syllable structure in a way that is irrelevant to the point under consideration, so I have given only examples of stems that end in a single consonant. Other data makes it clear that this is not a rule of epenthesis: *tate* 'father', *(dem) tatn* DATIVE.

**13.** This point is too obvious to be credited to Hale & Reiss or anyone else, for that matter. However, it seems to be ignored in OT arguments which suppose that the factorial typology, the set of possible ranking of a constraint set, should reflect *attested* languages. Obviously, the factorial typology must generate all attested patters, but it is clear that some may not be attested, for reasons that have nothing to do with phonology.

**14.** See Harnad 1987 for critical discussion of feature detector theory. Current work that attempts to incorporate phonetic description (acoustic parameters, trajectory of articulators, *etc.*) into the phonology represents an overly naive approach – one that essentially equates physical parameters with representational constants.

## References


Archangeli, D. & Pulleyblank, D. (1994). *Grounded Phonology*. Cambridge, MA: MIT Press.

Bhatia, T. & Kenstowicz, M. (1972). Nasalization in Hindhi: A reconsideration. *Papers in Linguistics, 5*, 202–212.

Bliese, L. F. (1981). A generative grammar of Afar. Summer Institute of Linguistics, University of Texas, Arlington.

Chomsky, N. (1986). *Knowledge of Language*. Westport, CT: Praeger.

Hale, M. (Forthcoming). *Theory and Method in Historical Linguistics.* Cambridge: Blackwell.

Hale, M. & Reiss, C. (2000a). Substance abuse and dysfunctionalism: current trends in phonology. *Linguistic Inquiry, 31*, 157–169.

Hale, M. & Reiss, C. (2000b). Phonology as cognition. In N. Burton-Roberts, Philip Carr and Gerry Docherty (Eds.), *Phonological Knowledge* (pp. 161–184). Oxford: Oxford University Press.

Halle, M. (1975). Confessio Grammatici. *Language, 51*, 525–535.

Halle, M., Vaux, B., & Wolfe, A. (2000). On feature spreading and the representation of place of articulation. *Linguistic Inquiry, 31*, 387–444.

Harnad, S. (Ed.) (1987). *Categorical Perception: The Groundwork of Cognition*. Cambridge: Cambridge University Press.

Jensen, J. T. (1977). *Yapese Reference Grammar*. Honolulu: University Press of Hawaii.

Keer, E. (1999). Geminates, the OCP and the Nature of CON. PhD thesis. Rutgers University, New Brunswick, New Jersey.

Lynch, J. (1978). A grammar of Lenakel. *Pacific Linguistics* Series B, 55. Australian National University, Canberra.

Marlett, S. & Stemberger, J. (1983). Empty consonants in Seri. *Linguistic Inquiry, 14*, 617–639.

McCarthy, J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry, 17*, 207–263.

Ni Chiosáin, M. & Padgett, J. (1993). Inherent V-Place. UC Santa Cruz Working Papers. Linguistics Research Center, University of California, Santa Cruz.

Odden, D. (1988). Antiantigemination and the OCP. *Linguistic Inquiry, 19*, 451–475.

Ohala, J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston and M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech.* Cambridge: Cambridge University Press.

Perlmutter, D. (1988). The split morphology hypothesis: evidence from Yiddish. In Michael Hammond and Michael Noonan (Eds.), *Theoretical Morphology*. San Diego: Academic Press.

Pylyshyn, Z. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science.* Cambridge, MA: MIT Press.

Quine, W. V. (1972). Methodological reflections on current linguistic theory. In G. Harman and D. Davidson (Eds.). *Semantics of Natural Language*. New York: Humanities Press.

Sapir, E. (1915). Notes on Judeo-German phonology. *The Jewish Quarterly Review*, n.s., 6, 231–266. Reprinted in *Selected Writings of Edward Sapir in Language, Culture, and Personality*. David G. Mandelbaum (Ed.). Berkeley: University of California Press.

Sherwood, D. (1986). *Maliseet-Passamaquoddy Verb Morphology.* Ottawa, Ontario: National Museums of Canada.

Taylor, S. (1969). Koya: an outline grammar, Gommu dialect. *University of California Publications in Linguistics* 54. Berkeley and Los Angeles: University of California Press.

Yip, M. (1988). The Obligatory Contour Principle and phonological rules: a loss of identity. *Linguistic Inquiry, 19*, 65–100.

# Contrast and asymmetries in inventories

B. Elan Dresher
University of Toronto

## 0. Introduction

There are a number of sources of asymmetry in phonology. One has to do with the different possibilities accorded to phonological heads as opposed to dependents (Dresher & van der Hulst 1998). In this paper I would like to focus on asymmetry that derives from the contrastive function of features. I assume that feature specifications are, at least to some degree, driven by considerations of contrast. A segment must be sufficiently specified so that it is differentiated from other segments in the inventory: the greater the number of contrasts, the more specifications will be required. Specifications that are required to draw a contrast between segments are *contrastive* with respect to those segments; specifications that are not contrastive are *redundant*. Phonologists have debated to what extent redundant specifications play a role in phonology. But the debate has necessarily been inconclusive, because of a lack of a clear notion as to what makes a specification contrastive or redundant in any given instance.

I will argue that in any inventory characterized by more than one distinctive feature it is necessary to specify the relative *scope* of each feature by means of a *contrastive hierarchy*. I will begin with some general observations about the logic of contrast. Then I will briefly look at the history of the notion of the contrastive hierarchy in phonological theory. Finally, I will illustrate the workings of the contrastive hierarchy with respect to synchronic and diachronic patterns in the vowel systems of the Manchu languages.

## 1. The logic of contrast

I'd like to begin with some general observations about the logic of contrast. I will employ non-linguistic items at first, so that it is clear that we are talking about contrast in a general way, and not about anything unique to phonology or morphology, or even language.

To begin, consider a simple inventory made up of the two items in (1).

(1) Inventory 1: Full specification

| | A. ■ | B. ◯ |
|---|---|---|
| Shape | square | round |
| Size | small | large |
| Colour | shaded | clear |

In (1) there are two contrasting figures. If these are the only items in the inventory, we may take them to be primitive elements, not further decomposable into constituent properties. Nevertheless, it is human nature to attempt to do just that, to try to identify characteristic distinguishing properties, or *features*. We observe that the two items differ on a number of dimensions:

a. Shape: Item A is square, B is a circle.
b. Size: Item A is smaller, B is larger.
c. Colour: Item A is shaded, B is clear.

Given just A and B, we have no way to tell whether all these features are essential to the identity of the item, or if any is more important than the others. It could be that one of these features is criterial, or *relevant* or *contrastive,* and the others are simply along for the ride, or *redundant*. It is also possible that some of these features form a gestalt that cannot be unpacked. Without some further information about how the members of this inventory behave, we can't tell.

Now let us look at Inventory 2, as shown in (2). Here we have three contrasting members: a small square (A), a large square (B), and a large circle (C). These members can be characterized by two features, shape and size. They are thus *distinctive* features, because they play a role in distinguishing between members.

(2) Inventory 2: Full specification

| | A.  | B.  | C.  |
|---|---|---|---|
| Shape | square | square | round |
| Size | small | large | large |

There is no doubt that each feature is distinctive somewhere in the inventory: shape alone distinguishes B from C, and size alone distinguishes A from B. But are both features *contrastive* with respect to each member of the inventory? In other words, do we need to specify both features for each member in order to uniquely distinguish it from the others? The answer is clearly no. In Inventory 2, A is the only member that is small, and C is the only member that is round. It follows that the feature [small] is sufficient to characterize A, and [round] immediately picks out C. Therefore, it ought to follow, so one might think, that a more economical specification would be as in (3).

(3) Inventory 2: Contrastive specification (incorrect!)

| | A.  | B.  | C.  |
|---|---|---|---|
| Shape | | square | round |
| Size | small | large | |

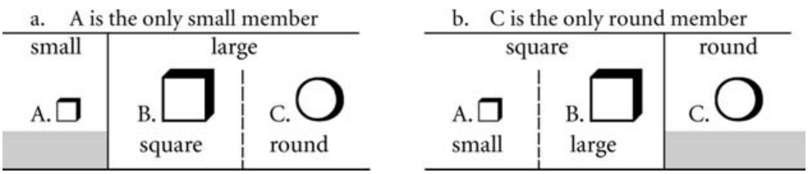
Let's say that a *contrastive specification* is one that uses only as many specifications as are required to distinguish each member from every other member. Is (3) a contrastive specification? Let's see if every member is minimally distinguished from every other one.

There are three pairwise comparisons to make. First, A and B are distinguished by size: A is [small] and B is [large]. B and C have a minimal contrast based on shape, [square] vs. [round].

But consider now A and C: the first is contrastively [small], and the second is contrastively [round]. Do these specifications allow us to draw a contrast between them? The answer is no. Their specifications look different, but they are not distinct. Indeed, the relation between A and C with respect to shape is like that between A and B: one is specified, the other is not. But A and B have the same shape. There is nothing in the chart in (3) that tells us that in this case the unspecified shape is not the same as the specified one. Similar considerations hold for size. We can, of course, write a rule that no specification for shape results in [square] and no specification for size results in [large]. Without applying at least one of these rules, we cannot tell if A and C are in contrast or not. But then we have failed in our attempt to represent all the relevant contrasts in the chart.

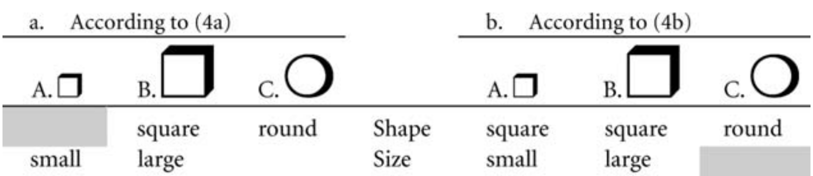
What has gone wrong here? In fact, the chart in (3) results from a misconstrual of our original observations about the inventory. When I wrote above that "A is the only member that is small," what I had in mind was that once we made a size contrast between A on one side ([small]) and B and C on the other ([large]), there is no need to further specify A for shape. The relevant contrasts can be pictured as in (4a). And when I wrote, "C is the only member that is round," I had in mind a picture such as (4b), where once C is specified [round] and A and B [square], there is no need to further specify C.

(4) Inventory 2: Two ways of viewing the contrasts




The diagram in (4a) corresponds to the feature specifications in (5a), and (4b) corresponds to (5b).

(5) Inventory 2: Two sets of contrastive specifications

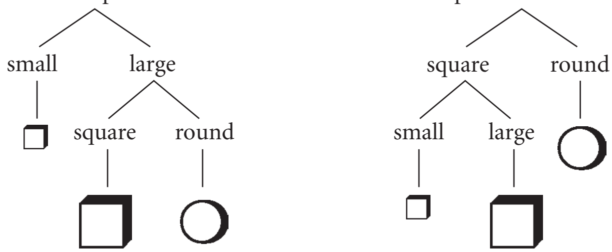



Diagrams (4a) and (4b) and their corresponding feature charts in (5a) and (5b) represent two distinct ways of "cutting up" the inventory. In (4a), the main contrast is size. We first divide up all the members of the inventory into two sets, one [large] and one [small]. Then we make a shape distinction in the [large] set. The [small] set has only one member, so no further contrasts in that set could be contrastive. Therefore, member A is contrastively specified for size, but only redundantly specified for shape. However, B and C are contrastively specified for both size and shape.

In (4b) the inventory is cut differently. Here, the first cut is by shape, and the second one is by size. The difference between (4a) and (4b) is in the *ordering,* or *scope,* of the features. In (4a), the size feature is ordered *before* the shape feature (i.e., has *wider* scope), and in (4b) size is ordered *after* shape (has *narrower* scope). Another way to represent these cuts is by tree diagrams, as in (6a) and (6b). The notation A > B means "A is ordered before B," or "A takes wider scope than B."

(6) Inventory 2: Tree diagrams of contrastive cuts

a. Size > Shape b. Shape > Size




Thus, the observations that A is the only small member and C is the only round member are correct, but in terms of contrastive force they derive from two different ways of cutting up the inventory. The ultra-minimal specification in (3) results from trying to put together two observations that derive from incompatible ways of cutting up the inventory. For this reason, it fails to adequately contrast A and C.

One reason that (3) may have looked correct is that, unlike (5a) and (5b), it contains no *logically redundant* features. We can define logical redundancy as in (7).

(7) Logical redundancy
If $\Phi$ is the set of feature specifications of a member, M, then the feature specification [F] is *logically redundant* iff it is predictable from the other specifications in $\Phi$.

A feature that is logically redundant is predictable from the other features. In Inventory 2, the feature [square] is predictable from [small], and [large] is predictable from [round]. However, logical redundancy is not sufficient to make a feature redundant *in any particular contrastive hierarchy*.

To see this very clearly, consider again Inventory 1, with only two contrasting members. Here, every feature is predictable given the other two. Indeed, any two features are predictable given the other one. Thus, all three features are logically redundant, but they can't all be omitted! Rather, logical redundancy indicates at most *potential* redundancy. Removing all logically redundant features often results in incoherence.

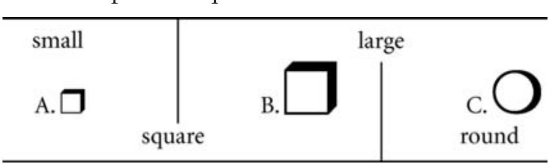
Another reason that we may not notice immediately that (3) fails to provide a set of adequate contrasts is that, when presented with a set of contrasting items, there is a tendency to concentrate on *minimal pairs*. A minimal pair is a pair of set members that differ from each other with respect to only one feature. It makes sense to focus on minimal pairs, because minimal pairs reveal the contrasting features in the purest way. It appears that many phonologists have implicitly or explicitly adopted a version of what we can call the Minimal Pairs Test.

(8) The Minimal Pairs Test (incorrect!)
To ensure that all members of an inventory are in contrast with each other, it is sufficient to specify contrasts between all potential minimal pairs.

The Minimal Pairs Test fails to ensure that all members of an inventory are in contrast. As with redundancy, minimal pairs with respect to full specification of distinctive features (*weak* minimal pairs) are only potentially minimal pairs given a particular contrastive hierarchy (*strong* minimal pairs). Thus, the large square, member B in Inventory 2, is weakly a minimal pair with both members A and C. But as the diagrams in (6) make clear, it forms a strong minimal pair with only one of these in any hierarchy.

There is a third logical possibility for how to characterize the contrasts in this inventory. Suppose that size and shape were *equal* in scope, so that none takes precedence over the other. Then we have to imagine the two ways of cutting up the inventory superimposed on each other, resulting in true *cross-classification*. This results in the full specification of (2), not the underspecification of (3). For in this approach, both features of every member are contrastive. In true cross-classification, we evaluate each feature *independently* from every other one. Looking *only* at shape, A and B are [square] and C is [round]. Looking *only* at size, A is [small] and B and C are [large]. On this model, member B is equidistant from A and C, as shown schematically in (9).

(9) Size = Shape: B is equidistant from A and C




What we conclude from the above discussion is that, when presented with an inventory and list of distinctive features, there is not necessarily a unique solution to what the contrastive specifications are. In any system of sufficient complexity, there will be a number of possible ways of structuring the system, depending on the ordering of relevant features. If ordering is completely free, then for *n* distinctive features there are *n*! possible orderings.

Not every ordering leads to a unique set of contrastive specifications. In general, the more symmetric an inventory is, the more the different feature orderings converge on the same specifications. That is, a symmetric inventory more closely approaches the appearance of cross-classification. The more asymmetric the inventory, the more dramatic the effect of different orderings. Of course, the apparent symmetry of an inventory is itself partially a function of the way the features are ordered.

## 2. The contrastive hierarchy via the Successive Division Algorithm

Our discussion of domains of contrast and the relative scope of features suggests that a natural way of determining contrast is by splitting the inventory by means of successive divisions, governed by an ordering of features. An algorithm corresponding to this idea, the Successive Division Algorithm (SDA), is given in (10):

(10) Successive Division Algorithm (SDA)

- a. In the initial state, all tokens in inventory, I, are assumed to be variants of a single member. Set I = S, the set of all members.
- b. i. If S is found to have more than one member, proceed to (c).
  - ii. Otherwise, stop. If a member, M, has not been designated contrastive with respect to a feature, G, then G is *redundant* for M.
- c. Select a new *n*–ary feature, F, from the set of distinctive features.[1] F splits members of the input set, S, into *n* sets, $F_1 - F_n$, depending on what value of F is true of each member of S.
- d. i. If all but one of $F_1 - F_n$ is empty, then loop back to (c).[2]
  - ii. Otherwise, F is *contrastive* for all members of S.
- e. For each set $F_i$, loop back to (b), replacing S by $F_i$.

The algorithm in (10) is a very general formulation for defining contrastiveness and redundancy for members of an inventory. It can apply to *n*–ary features and it does not tie contrastiveness and redundancy to how a member is specified. In practice, though, the most natural assumption to make is that all members that are contrastive for a feature, F, receive specifications for F, and members that are redundant do not.

## 3. Binary (equipollent) vs. unary (privative) features

Another source of asymmetry in phonology comes from the nature of the features themselves. So far I have assumed binary features where the two values are equal (*equipollent* features). But this may not be what linguistic features are like. It could be that one value is *marked* and the other is *unmarked*. We could take markedness literally, so that only the marked value is marked, and the unmarked value is simply the absence of the marked one. This is a *privative* feature system.

To make an equipollent contrast between round and square, it is enough to write [round] ~ [square], or [+round] ~ [–round], or [–square] ~ [+square]. To make a privative contrast, we have to decide which is the marked feature. For the sake of discussion, let us set up the following privative features as in (11).

(11) Privative features for size, shape, and colour

| | Size | Shape | Colour |
|---|---|---|---|
| marked | large | round | shaded |
| unmarked | (small) | (square) | (clear) |

In (11) the features in parentheses are defaults that do not actually appear in contrastive specifications. They are there only to remind us what the absence of a value entails. The only actual marks are the marked values.

Fully specifying all distinctive features for Inventory 2 now results in (12).

(12) Inventory 2: Full specification (privative features)

| | A.  | B.  | C.  |
|---|---|---|---|
| Shape | | | round |
| Size | | large | large |

Note that the specifications of A and B are already minimal, so only the specifications of C can be affected by ordering. If size is the first feature, then we observed (6a) that A is the only member that is small. Using equipollent features, we marked A [small] and B and C [large]. But now the former is not a specification, so we leave A unmarked and mark B and C as [large]. The next contrast is between B and C on the basis of shape, and we mark C as [round]. Now we have the same specifications as in (12). Proceeding in the other order, we first mark C [round] in contrast to A and B, which are unmarked. We then draw a contrast between A and B by marking B [large]. This time, C is specified only as [round].

The effect of feature ordering is reduced with privative features as compared with equipollent features. For similar reasons, the relationship between contrastiveness and specification is more transparent with equipollent than with privative features. This is because in an equipollent system, if a member, M, is specified for a feature, F, then F is contrastive for M. Conversely, if M is unspecified for F, then F is redundant for M. In a privative system, specification indicates contrastiveness, as in an equipollent system. However, if M is unspecified for F, then *either* i) F is redundant for M, *or* ii) M is unmarked for F (or both (i) and (ii) are true).

## 4. A very brief history of the contrastive hierarchy in phonology

Since Saussure it has been commonly stated that contrast is central to linguistic theory, but there has been relatively little discussion of how contrastiveness is to be assigned in general. With a few notable exceptions, there is almost nothing in the phonological literature on the notion of a contrastive hierarchy, much less that it is central to establishing contrast. Nevertheless, a hierarchy has been implicit in at least a partial way in the practice of phonologists.

One place to see this is in the way tables of segmental inventories are arranged. Compare, for example, the tables of Siglitun (13) and Kolokuma Ịjọ (14). I present them as they are given in the sources (with some changes to the phonetic symbols but not to the arrangement). Note in particular the different placements of /l/ and /j/ in these charts. The chart of Ịjọ expresses a hierarchy whereby the feature [continuant] has wider scope than such features as [sonorant] and [voiced], and [lateral] has wider scope than [nasal]. The Siglitun chart is not as overtly hierarchical, but it is clear that the feature [lateral] has very narrow scope, confined to making distinctions among apicals, whereas [nasal] is higher in the hierarchy. Apart from the nasals, the other sonorants are not set apart in Siglitun, suggesting that the feature [sonorant] is lower in the hierarchy than in Ịjọ.

(13) Siglitun consonants (Dorais 1990: 70)[3]

| | Bilabial | Apical | | Velar | Uvular |
|---|---|---|---|---|---|
| Stops | p | t | | k | q |
| Voiced fricatives | v | l | j | ɣ | R |
| Voiceless fricatives | | ÿ | s | | |
| Nasals | m | n | | ŋ | |

(14) Consonant phonemes of Kolokuma Ịjọ (Williamson 1965)[4]

| | Plosive | | Continuant | | | | |
|---|---|---|---|---|---|---|---|
| | | | Fricative | | Sonorant | | |
| | Vl. | Vd. | Vl. | Vd. | Non-lateral | | Lateral |
| | | | | | Oral | Nasal | |
| Labial | p | b | f | v | w | m | |
| Alveolar | t | d | s | z | r | n | l |
| Back | k | g | (h) | (ɣ) | j | ŋ | |
| Labio-velar | kp | gb | | | | | |

Trubetzkoy (1969, completed in 1939) does not explicitly mention the contrastive hierarchy, nor does he consistently follow it. However, many of his examples presuppose a notion of relative scope. In his discussion of the Polabian vowel system, for example, he explicitly refers to a hierarchy of contrasts. He writes (1969: 102–103) that a "certain hierarchy existed" whereby the back ∼ front contrast is higher than the rounded ∼ unrounded one, the latter being a subclassification of the front vowels. He also proposes (1969: 102) that "the maximally open vowel a. . . stood outside the classes of timbre" (i.e., the front ∼ back classes). This can only mean that /a/ is split off from the rest of the vowels first. The vowel system, according to Trubetzkoy's contrastive distinctions, is given in (15).

(15) Polabian (Trubetzkoy 1969: 102–103): Back > Rounded

| Front | | Back |
|---|---|---|
| (Unrounded) | Rounded | |
| i | ü | u |
| ê | ö | o |
| e | | ɑ |
| | a | |

As far as I know, the first person to advocate the contrastive hierarchy was Roman Jakobson. Even he employed it inconsistently. The hierarchy was given a prominent place in Jakobson and Halle's *Fundamentals of Language* (1956). Jakobson and Halle (1956) refer to this hierarchy as the "dichotomous scale," and adduce "several weighty arguments" in support of this hierarchical approach to feature specification. First, they claim (1956: 47) that such a system is an "optimal code" for speech participants who have to encode and decode messages.

Their second argument involves language acquisition. They suggest that distinctive features are necessarily binary because of the way they are acquired, through a series of "binary fissions". They propose (1956: 41) that the order of these contrastive splits is partially fixed, thereby allowing for certain developmental sequences and ruling out others. They propose the sequence in (16) for oral resonance (primary and secondary place) features. The decimals indicate precedence relations: if one decimal sequence is entirely contained in another sequence, then the contrast corresponding to the former must precede the acquisition of the latter contrast. No implicational relations exist between features represented by an equal number of decimals. Thus, the chart in (16) is equivalent to a branching tree diagram, or a partially ordered lattice, of the sort that follows from the Successive Division Algorithm.

(16) Predicted acquisition sequences (Jakobson & Halle 1956: 41)

| | |
|---|---|
| Consonants: dental vs. labial | 0.1 |
| Vowels: narrow vs. wide | 0.11 |
| Narrow vowels: palatal vs. velar | 0.111 |
| Wide vowels: palatal vs. velar | 0.1111 |
| Narrow palatal vowels: rounded vs. unrounded | 0.1112 |
| Wide palatal vowels: rounded vs. unrounded | 0.11121 |
| Velar vowels: unrounded vs. rounded | 0.1113 |
| Consonants: velopalatal vs. labial and dental | 0.112 |
| Consonants: palatal vs. velar | 0.1121 |
| Consonants: rounded vs. unrounded or pharyngealized vs. non-pharyngealized | 0.1122 |
| Consonants: palatalized vs. non-palatalized | 0.1123 |

The notion of a developmental sequence of phoneme acquisition did take hold in the field of child language. This literature, however, remained isolated from the mainstream of theoretical phonology.

Halle employed the contrastive hierarchy in *The Sound Pattern of Russian* (1959), though he did not emphasize its empirical nature. Also, he did not consider whether the representations he derived were the ones that best fit the patterning of the phonology, nor did he discuss the reasons for his choice of hierarchy. Rather, he argued more generally that the hierarchy is required to ensure that segments are distinct, that is, properly in contrast. Halle (1959: 32) formulates a requirement that phonemes meet the Distinctness Condition (17).

(17) Distinctness of phonemes (Halle 1959: 32)
Segment-type {A} will be said to be *different from* segment-type {B}, if and only if at least one feature which is phonemic in both, has a different value in {A} than in {B}; i.e., plus in the former and minus in the latter, or vice versa.

He gives the examples in (18).

(18) Examples of distinctness and non-distinctness (Halle 1959: 32)

a. {A} is not "different from" {C}

| | {A} | {B} | {C} |
|---|---|---|---|
| Feature 1 | + | – | + |
| Feature 2 | 0 | + | – |

b. All three segment-types are "different"

| | {A} | {B} | {C} |
|---|---|---|---|
| Feature 1 | + | – | – |
| Feature 2 | 0 | + | – |

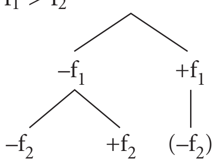
(17) and (18) are essentially the argument I gave at the outset. For two items to be in contrast, they must be distinct in at least one feature. The lack of a specification does not make a segment distinct in an equipollent system, which Halle (1959) assumes. Halle (1959) argues that only a "branching diagram", i.e., a contrastive hierarchy, can ensure that the Distinctness Condition is met.

The use of "branching diagrams" was challenged by Stanley (1967), as part of his famous article arguing against unspecified features in the phonology. With respect to branching diagrams, he argued first that the Distinctness Condition is both too strong and too weak. It is too weak because it does not prevent "specious simplifications" of rules obtained by improper use of zeros. This is of course part of Stanley's general argument against zeros in the phonology, and I cannot address that here. The Distinctness Condition is too strong, in Stanley's view, because it prevents leaving out redundant features when feature values of a higher-order feature are predictable from the values of lower-order features (what we can call "underspecification from below").

For example, suppose we have a feature hierarchy as in (19a). If a segment is [$+f_1$], then it must be [$-f_2$]; since this specification is not contrastive, our algorithm would omit it. For clarity, I indicate it here in parentheses. This value of $f_2$ can be filled in by the redundancy rule (19bi). However, it is now also true that if a segment is [$+f_2$] then it must also be [$-f_1$]; Stanley argues that this fact, expressed by rule (19b-ii), should allow us to underspecify two segments that contrast with respect to $f_1$ as in (19c).

(19) Distinctness Condition and redundancies from below (Stanley 1967)

a. $f_1 > f_2$




b. Redundancy rules

i. [$+f_1$] $\rightarrow$ [$-f_2$] ii. [$+f_2$] $\rightarrow$ [$-f_1$]

c. Lexical entries violating the DC

| | A | B |
|---|---|---|
| $f_1$ | + | |
| $f_2$ | | + |

d. Full specification

| | A | B |
|---|---|---|
| $f_1$ | + | – |
| $f_2$ | – | + |

Such situations occur frequently in languages. A typical case is a language where all sonorants are voiced; then, a specification of [–voiced] necessarily implies [–sonorant]. Stanley argues that the lexical entries in (c) should be permitted, because the redundancy rules in (b) are legal and lead to the two segments being completely distinguished (d).

Stanley's argument would go through if it were indeed a requirement of the theory that all logical redundancies be expressed by zero specifications. But we have seen before that it is neither necessary nor possible for a theory to express all logical redundancies. If we take the contrastive hierarchy as the fundamental way to express contrasts in a system, then a certain amount of underspecification will follow from it. However, underspecification, on this view, is a consequence of establishing contrasts, not an end in itself. We would thus rule out specification from below as in rule (19bii): lower-order features cannot supply values for higher-order features.

Stanley's second argument against the branching trees is that it is "somewhat strange" to capture the feature hierarchy in a branching diagram because many different such diagrams can be constructed for a given set of phonemes. Stanley is correct that one must have evidence to support any proposed hierarchies. He even suggests (1967:408) where such evidence could possibly be found, "perhaps...in terms of the different ways in which different features behave in the P[honological] rules or the M[orpheme]S[tructure] rules." At the time he did not think he had such evidence, and indeed it was not (and is still not!) entirely clear what would constitute such evidence. However, we have seen that from the beginnings of distinctive feature theory phonologists, whether in theory or in practice, have found reasons to assign at least partial feature hierarchies.

In sum, Stanley's arguments against the branching trees, and hence against the contrastive hierarchy, rest mainly on logical, not empirical, criteria. The logical points are correct as far as they go, however his main empirical assumptions have become less relevant as phonological theory developed. Nevertheless, his arguments carried the day, and branching trees disappeared from generative phonology for the rest of the century.

However, the intuition that, as Stanley put it (408), "there is obviously some kind of hierarchical relationship among the features which must somehow be captured in the theory", continued to haunt generative phonological theory, and took a number of different forms. Three of them coexisted uneasily at the same time at the heart of mainstream generative phonology: markedness theory, underspecification theory, and feature geometry. All bear interesting, though seldom discussed, affinities with the contrastive hierarchy.

## 5. Contrast in Manchu vowel systems

I will conclude with a more recent example of an analysis that makes use of the contrastive hierarchy (see Dresher & Zhang 2000 for a more detailed account). Zhang (1996) argues that contrastive specification following the contrastive hierarchy provides an illuminating account of the vowel system of Written Manchu, as well as of the evolution of the later Manchu languages.

Written Manchu has the vowel system in (20).

(20) Written Manchu (Zhang 1996)

| | | |
|---|---|---|
| i | | u |
| | | ʊ |
| | ə | |
| | | ɔ |
| | a | |

Vowels are divided into co-occurring sets based on ATR harmony (Zhang 1996; Zhang & Dresher 2000): /ə/ and /u/ spread ATR to /ʊ/ and /a/. /i/ is neutral: it occurs with both sets of vowels. Even though it is phonetically ATR, it does not trigger ATR harmony. Why is that? The answer is that ATR has narrow scope in Written Manchu. If the contrastive hierarchy is as in (21), then /i/ is contrastively Coronal, but does not receive a specification for ATR.

(21) Written Manchu vowel contrasts: Low > Coronal > Labial> ATR

| | | | |
|---|---|---|---|
| Coronal | i | ATR u<br>ʊ | |
| Low | | ə ATR<br>a | ɔ Labial |

On this contrastive hierarchy, ATR is a redundant feature for /i/, filled in by the rule in (22).

(22) /i/ becomes ATR phonetically

$$\begin{array}{l} \quad \mathrm{V} \\ [\mathrm{Coronal}] \rightarrow \mathrm{ATR} \end{array}$$

Once /i/ is specified Coronal, there is no need to specify the back high vowels as Labial. In the Low region, /a/ is too far back to be Coronal, so /ɔ/ must be Labial. These specifications accord with the patterning of labial harmony, which is triggered only by the Low round vowel. /u/ and /ʊ/ do not trigger labial harmony (see Zhang and Dresher 1996 for further details on labial harmony). The contrastive vowel features of Written Manchu are thus as in (23).

(23) Feature matrix for Written Manchu vowels

| Phoneme | /i/ | /u/ | /ʊ/ | /ə/ | /a/ | /ɔ/ |
|---|---|---|---|---|---|---|
| Low | | | | √ | √ | √ |
| Coronal | √ | | | | | |
| Labial | | | | | | √ |
| ATR | | √ | | √ | | |

Diachronic developments support this analysis. Spoken Manchu, a descendant of Written Manchu (or of a language closely related to Written Manchu) has the vowel system shown in (24).

(24) Spoken Manchu (based on Zhao 1989; Ji et al. 1989)

| | | | |
|---|---|---|---|
| i | y | ə | u |
| ɛ | | a | ɔ |

The differences between the vowel systems of Written Manchu and Spoken Manchu are summarized in (25).

(25) Differences between WM and SM vowel systems

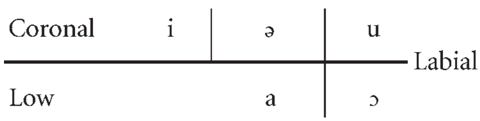
a. Spoken Manchu no longer has a contrast between /u/ and /ʊ/.

b. In Spoken Manchu /ə/ is a non-low vowel, in Written Manchu it is a low ATR counterpart to /a/.

c. Spoken Manchu has added coronal phonemes /y/ and /ɛ/.

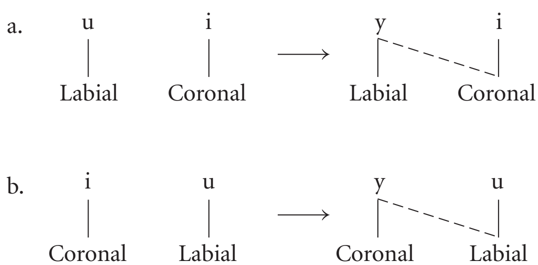
These differences are all connected. It appears that Spoken Manchu first lost the distinction between /u/ and /ʊ/. This distinction was already neutralized at the surface in many environments in Written Manchu, and in Spoken Manchu the neutralization was completed. Consequently, the ATR contrast would have rested only on /a/ and /ə/. But this contrast could now be reinterpreted as one of height, so that /ə/, of ambiguous height, was reclassified as a non-low vowel, as in (26).

(26) Spoken Manchu after loss of /ʊ/




Now an additional contrast has to be made among the non-low vowels: /i/ is Coronal, and /u/ becomes Labial to be distinguished from /ə/. Evidence that /u/ has taken on a Labial specification is provided by the development of a new phoneme /y/, a front rounded vowel created from combining /i/ and /u/ (27). The rounding could only come from /u/.

(27) Creation of /y/ from /u/ - /i/ and /i/ - /u/




Further evidence comes from Xibe, another descendent of Written Manchu, whose vowel system is shown in (28).

(28) Xibe (Sibo) (based on Li & Zhong 1986)

| Coronal | | | | |
|---|---|---|---|---|
| | Labial | | Labial | |
| i | y | ə | u | |
| ɛ | œ | a | ɔ | Low |

As in Spoken Manchu, Xibe /u/ has acquired a Labial specification. Further support for this analysis comes from the operation of labial harmony in Xibe, which is triggered by /u/. Recall that /u/ in Written Manchu did not trigger labial harmony.

## Conclusion

If the approach to contrastive specification sketched above is correct, it remains to be seen to what extent particular contrastive hierarchies are universal, and to what extent they can vary. This topic is the subject of ongoing work in the project on Contrast and Complexity in Phonology at the University of Toronto (http://www.chass.utoronto.ca/∼contrast/). Some relevant references are: Avery (1996), Avery and Rice (1989), Balcaen (1998), Causley (1999), Dresher (1998a, b), Dresher and Rice (1993), Dresher, Piggott, and Rice (1994), Dyck (1995), Ghini (2001), Hall (1998), Rice (1993, 1995, 1997), Rice and Avery (1995), Walker (1993), Wu (1994), Zhang (1996), and Zhou (1999).

## Acknowledgements

I would like to thank the members of the project on Contrast and Complexity in Phonology in the Department of Linguistics at the University of Toronto for many kinds of help over the years. I am especially indebted to Xi Zhang for the analysis of Manchu. This research was supported in part by grants 410-96-0842 and 410-99-1309 from the Social Sciences and Humanities Research Council of Canada.

## Notes

**1.** I assume that the set of relevant distinctive features for a particular domain is given by some theory of that domain. By "new" feature I mean one that has not already been tried. Thus, the value of F changes every time this step reapplies (I assume some mechanism for keeping track of which features have already been tried, but do not specify it here).

**2.** That is, if all members of S have the same value of F, then F is not contrastive in this set.

**3.** I have simplified Dorais's *j/dj* and *s/ch* to *j* and *s*, respectively. As he makes clear, these are variants of single phonemes. Dorais does not usually indicate variants in his charts, and in related dialects in which j has similar variants he lists only *j*. Therefore, I keep to the usual practice of representing a phoneme by one symbol.

**4.** I substitute *j* for Williamson's *y*. Williamson notes that Back = palatal, velar or glottal, Vl. = voiceless, and Vd. = voiced. Williamson mentions that some speakers have a marginal phoneme /ɣ/, but she omits it from the table. I have added it because it appears to be no less marginal than /h/, which is included.

## References


Avery, P. (1996). The representation of voicing contrasts. Doctoral dissertation, University of Toronto.

Avery, P. & Rice, K. (1989). Segment structure and coronal underspecification. *Phonology, 6*, 179–200.

Balcaen, M. J. (1998). Neutralization and enhancement: Vowel harmony in Aka. Ms., University of Toronto.

Causley, T. (1999). Complexity and markedness in Optimality Theory. Doctoral dissertation, University of Toronto.

Dorais, L.-J. (1990). *Inuit Uqausiqatigiit: Inuit Languages and Dialects*. Iqaluit: Arctic College – Nunatta Campus.

Dresher, B. E. (1998a). Child phonology, learnability, and phonological theory. In T. Bhatia and W. C. Ritchie (Eds.), *Handbook of Language Acquisition* (pp. 299–346). New York: Academic Press.

Dresher, B. E. (1998b). On contrast and redundancy. Paper presented at the annual meeting of the Canadian Linguistic Association, Ottawa. Ms., University of Toronto.

Dresher, B. E. & van der Hulst, H. (1998). Head-dependent asymmetries in phonology: Complexity and visibility. *Phonology, 15*, 317–352.

Dresher, B. E. & Rice, K. (1993). Complexity in phonological representations. *Toronto Working Papers In Linguistics, 12*, i–vi.

Dresher, B. E. & Zhang, X. (2000). Contrast in Manchu vowel systems. Paper presented at the 1st International Conference on Manchu-Tungus Studies (ICMTS 1), University of Bonn, August 28 – September 1, 2000.

Dresher, B. E., Piggott, G. & Rice, K. (1994). Contrast in phonology: Overview. *Toronto Working Papers In Linguistics, 13*, iii–xvii.

Dyck, C. (1995). Constraining the phonology-phonetics interface, with exemplification from Spanish and Italian dialects. Doctoral dissertation, University of Toronto.

Ghini, M. (2001). *Asymmetries in the Phonology of Miogliola*. Berlin: Mouton de Gruyter.

Hall, D. C. (1998). Contrastive specification for voicing in Czech. Paper presented at the MOT Phonology Workshop, University of Ottawa. Ms., University of Toronto.

Halle, M. (1959). The *Sound Pattern of Russian: A Linguistic and Acoustical Investigation*. The Hague: Mouton. Second printing, 1971.

Jakobson, R. & Halle, M. (1956). *Fundamentals of Language*. The Hague: Mouton.

Ji, Y, Zhao, Z. & Bai, L. (1989). *Xiandai Manyu Babaiju* [*Eight Hundred Sentences of the Modern Manchu*]. Beijing: Zhongyang Minzu Xueyuan Chubanshe.

Li, S. & Zhong, Q. (1986). *Xiboyu Jianzhi* [*A Brief Introduction to the Xibe Language*]. Beijing: Minzu Chubanshe.

Rice, K. (1993). A reexamination of the feature [sonorant]: The status of 'sonorant obstruents'. *Language, 69*, 308–344.

Rice, K. (1995). On vowel place features. *Toronto Working Papers In Linguistics, 14*, 73–116.

Rice, K. (1997). Japanese NC clusters and the redundancy of postnasal voicing. *Linguistic Inquiry, 28*, 541–551.

Rice, K. & Avery, P. (1995). Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In John Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 23–42). Hillsdale, NJ: Lawrence Erlbaum.

Stanley, R. (1967). Redundancy rules in phonology. *Language, 43*, 393–436.

Trubetzkoy, N. (1969). *Principles of Phonology*. Translated from *Grundzüge der Phonologie* (Göttingen: Vandenhoek & Ruprecht 1958; third edition, 1962) by C. A. M. Baltaxe. Berkeley: University of California Press.

Walker, R. (1993). Contrastive specification and vowel harmony mechanisms in the Altaic languages. Forum paper (M.A. thesis), University of Toronto.

Williamson, K. (1965). *A Grammar of the Kolokuma Dialect of Ijo*. Cambridge: Cambridge University Press.

Wu, Y. (1994). Mandarin segmental phonology. Doctoral dissertation, University of Toronto.

Zhang, X. (1996). Vowel systems of the Manchu-Tungus languages of China. Doctoral dissertation, University of Toronto.

Zhang, X. & Dresher, B. E. (1996). Labial harmony in Written Manchu. *Saksaha*: *A Review of Manchu Studies, 1*, 13–24.

Zhang, X. & Dresher, B. E. (2000). Tongue root harmony in Written Manchu. Paper presented at the 1st International Conference on Manchu-Tungus Studies (ICMTS 1), University of Bonn, August 28 – September 1, 2000.

Zhao, J. (1989). *Xiandai Manyu Yanjiu* [*Studies on the Modern Manchu Language*]. Beijing: Minzu Chubanshe.

Zhou, H. (1999). Vowel systems in Mandarin languages. Doctoral dissertation, University of Toronto.

# The acquisition of constructions with reflexive clitics in Polish*

María Luisa Rivero and Magdalena Golędzinowska
University of Ottawa

*Of course the data of performance have long ago been pronounced (Chomsky 1964) an inadequate base for a grammar that attempts to represent competence or knowledge. I agree that it always is but I venture to say that not many people know how much can be milked from mere performance in the case of small children.*

(Roger Brown, *A First Language: The Early Stages*, page 56)

## 1. Introduction

This paper has two aims; one is empirical and the other theoretical. The empirical aim is a description of the acquisition stages in Polish of constructions with reflexive clitics, and with modal *wolno* 'allow'. The theoretical aim is to argue that acquisition data on reflexive clitics and modal *wolno* provide new support for the Maturation Hypothesis (Babyonyshev et al. 2001; Borer & Wexler 1987, 1992). That is, UG is not available in its adult form in all stages of acquisition, but matures into the adult grammar. The set of representations allowed in proto-UG is a subset of representations allowed by UG. During grammatical development, UG principles are gradually reformulated, remaining a subset of adult grammar at all developmental stages.

Opponents of the Maturation Hypothesis most often present the argument known as the "poverty of stimulus": children master the grammar of the ambient language despite highly fragmented input and lack of explicit instruction (Pinker 1984, 1989; Hyams 1987, among others). The argument we set out to support is the opposite, and following Babyonyshev et al. (2001), it may be termed to arise from the observation of "abundance of the stimulus." More specifically, the Maturation Hypothesis receives support from the apparent delay in children's linguistic performance in the face of ample positive input. "The explanation," Babyonyshev at al. claim, "lies in the hypothesis that the biology that supports the relevant knowledge is not available until a comparatively late stage in child development." This hypothesis is the hereafter supported Maturation, whose main postulate states that "specific details of linguistic knowledge are biologically determined" (Babyonyshev et al. 2001).

Our proposals are based on the naturalistic output in the CHILDES database (MacWhinney 2000; Weist et al. 1984, 1986) of five children (**I**nka, **W**awrzon, **K**ubuś, **B**artosz, and **M**arta) and input from their caretakers. We did an exhaustive study of two sets of constructions: those with the reflexive clitic *się*, and those with modal *wolno* 'allow'. Using CLAN software, a word search was conducted on the corpora of transcribed conversations between the children and their caretakers. All the constructions with **się** and *wolno,* and with obligatory *się* and *wolno* omitted were included in the analysis. **W**, **K**, **B**, and **M**'s utterances were morphemically glossed and translated into colloquial English in the original transcriptions. **I**'s files did not contain any glosses; therefore, morphemic glosses and translations made by Magdalena Golędzinowska, a native speaker of Polish, are used.

The first group of investigated constructions is with the reflexive clitic *się* in functions listed in section 2, including unaccusatives, (1), and impersonals, (2).

(1) a. *Przewróciłem się.* **B** (1;8)
fall$_{\text{PAST.1SG}}$ REFL
'I fell.'

b. *Przewróciła się.* **M** (1;8)
fall$_{\text{PST.3SG.FEM}}$ REFL
'(The doll) fell.'

c. *Zmęczyłem się.* **B** (1;11)
tire$_{\text{PST.1SG.MASC}}$ REFL
'I got tired.'

d. *Okno się otworzyło.* **K** (2;2)
window$_{\text{NOM. NEU}}$ REFL open$_{\text{PAST.NEU}}$
'The window opened.'

(2) *Do prądu się podłącza prostownik.* **W** (2;9)
to current REFL plug+in$_{\text{PRES.3SG}}$ rectifier$_{\text{ACC.SG}}$
'One plugs the rectifier into the current.'

*Wolno* is an invariable modal with two syntactic arguments in adult grammar in brackets in (3a) – infinitive as theme of (non)-permission, and Dative as goal. The "impersonal" pattern without expressed goal is as in (3b).

(3) a. [*Mnie*] *nie* <u>*wolno*</u> [*tak daleko iść*] *prawda*? **I** (3;2)
[$\mathrm{I}_{\mathrm{DAT}}$] Neg allow [such $\mathrm{far}_{\mathrm{ADV}}$ $\mathrm{walk}_{\mathrm{INF}}$], $\mathrm{right}_{\mathrm{TAG}}$
'I am not allowed to walk that far, right?'

b. *Tatusiu*, *nie* <u>*wolno*</u> [*bić mamusi*]. **I** (1;9)
$\mathrm{daddy}_{\mathrm{VOC}}$ Neg allow $\mathrm{hit}_{\mathrm{INF}}$ $\mathrm{mommy}_{\mathrm{GEN}}$
'Daddy, {it, one} is not allowed to hit mommy.'

We propose that the acquisition of the two construction types supports Maturation, proceeding in three stages. Table 1 indicates the three stages for each child in clitic (*się*) constructions. In the case of **I**, the stages of *się* development chronologically coincide with those of *wolno*.

Stage I is marked by systematic omission. The child utters **Okno otworzyło* instead of the adult (1b) *Okno <u>się</u> otworzyło*, and lacks *wolno.* In Stage II, the required *się* emerges most prominently in unaccusatives. Some adult uses including impersonals of type (2) are absent, omissions continue, and overgeneralizations based on the early uses appear. In Stage III, a variety of clitic constructions appear, including impersonals, and *wolno* with an expressed goal matures as well.

We interpret the stages as follows. The first illustrates omission of functional categories (clitic *się* and modal *wolno*) in early grammar. The second shows the fitting of the clitic and of *wolno* into two distinct areas of the grammar guided by the principle Borer and Wexler dub the U(nique) E(xternal) A(rgument) P(roto)-P(rinciple) (1992: 31). Under this proto-principle, every predicate associates with a unique external argument regardless of verb type, and every external argument associates with a unique predicate. We argue that in Polish, the UEAPP forces the clitic to be a predicate modifier, as it requires that each predicate establish a predicative relation with a unique sister argument. In other words, the V denotes a property applied to a unique argu-

**Table 1.** Developmental stages in reflexive clitic constructions

| Child | | Age | Stages |
|---|---|---|---|
| Inka | (**I**) | 0;10 - 3;6 (...) | I, II, III |
| Wawrzon | (**W**) | 2;2 - 3;2 | I, II, III |
| Kubuś | (**K**) | 2;1 - 2;6 | II, III |
| Bartosz | (**B**) | 1;7 - 1;11 | II |
| Marta | (**M**) | 1;7 - 1;10 | II |

ment, and the clitic is like a syntactic adverb that specifies this property, or is a function from a property into a property: [*się* (P (opened))] (window).

During this stage (and likely into adulthood), unaccusativity and unergativity are partially similar properties applying to a unique syntactic argument. This parallelism is empirically supported by the unaccusative and unergative constructions with reflexive clitics in Section 2. On this view, unaccusatives of type (1) fit the UEAPP, so it does not come as a surprise that they can be found with an adult-like clitic as early as 1;7, and count among the first patterns with an emerging *się* in all five children. Similarly, the argument structure of unergative verbs, which appears in the proto-principle stage with a reflexive clitic, is compatible with the UEAPP.

Interestingly, not all clitic constructions develop as early. Impersonal *się* constructions, ordinary in the speech of the children's caretakers, are absent in **B** and **M,** and appear in adult form in **I**, **K** and **W** only after the age of 2;4. Why should they be delayed? In the Polish literature, impersonals are "active" sentences, and a recent version of this idea is that they have two syntactic arguments (Rivero 1999). On this view, the Acc(usative) NP in (2) is the internal argument, and the clitic signals the external argument. If this analysis is correct, Polish *się*-impersonals mature late because they are incompatible with the UEAPP, which rules them out in Stage II.

In Stage III the UEAPP is modified, allowing for the development of the late clitic constructions such as impersonals as well as **wolno**-personals with goal expressed as in (3a).

In sum, Polish unaccusatives with clitics develop very early, and impersonals with clitics develop later, and in this paper this disparity is assigned to syntax, the UEAPP, and Maturation.

As to *wolno*, this modal emerges in Stage II in the one-argument "impersonal" version without Dative goal, so our proposal is that it is also regulated by the UEAPP.

Inka, the child recorded in all clitic stages who also offers very rich data on *wolno*, proves particularly interesting for Maturation. During Stage II, **I** lacks *się*-impersonals and displays *wolno*-impersonals of type (3b). These two types of impersonals denote indeterminate humans, so the contrast is not due to semantic difficulties with impersonality, and here we derive it from the UEAPP, which blocks clitic impersonals but not *wolno*-impersonals. This is because clitic impersonal constructions require two arguments in the syntax, forbidden by the UEAPP, whereas *wolno*-impersonals are well formed with only one argument. As the principle matures in Stage III, **I**'s knowledge of the two areas of grammar changes, and two seemingly different constructions that display a

**Table 2. Stages of grammatical development in Inka**

| Stage | Age |
|---|---|
| Stage I | 0;10 – 1;9/1;10 |
| Stage II | 1;10 – 2;5 |
| Stage III | 2;5 – 3;6 (and later) |

richer syntactic argument structure mature: *się*-impersonals for indeterminate humans, and *wolno*-personals for determinate animates expressed as goal.

In their study of Russian unaccusatives, Babyonyshev et al. (2001) propose the A-C(hain) D(eficit) H(ypothesis), which states that A-chain formation is not available in the early stages of grammatical development. Although children do utter constructions analyzed to contain A-chains in the adult grammar, Babyonyshev et al. suggest that they assign representations without A-chains. The ACDH does not play a major role in our discussion for two reasons. One, we do not assume ordinary NP movement for Polish clitic constructions (some contain clitic chains). Two, development restrictions on *wolno* derive from the UEAPP, a precondition to the ACDH, not from movement or A-chains. However, A-chains could delay some constructions that develop after the period considered here, including unaccusative impersonals as in fn 4. Table 2 lists the ages corresponding to the three stages of grammatical development in Inka.

Section 2 lists the functions of the Polish clitic (Kubiński 1987; Dziwirek 1994; Rivero 1999 for earlier references and discussion). Section 3 examines the acquisition stages for clitic constructions, and Section 4 - for constructions with *wolno*.

## 2. Functions of the Polish reflexive clitic

This section lists the adult grammar functions of the reflexive clitic in their order of development in the CHILDES database, and sketches analyses for them. As stated, major classes that mature early in all children are unaccusatives including inchoatives, (1) and (4), and inherents as in (6). Word order is not relevant.

(4) a. *Chowa* *się* *dzidziuś.* **B** (1;7)
hide$_{\text{PRES.3SG}}$ REFL baby$_{\text{NOM}}$
'The baby is hiding.'

b. *Czemu kręcą* <u>*się*</u> *kółeczka*? **M** (1;9)
why turn $_{\text{PRES.3PL}}$ REFL little+wheels$_{\text{NOM}}$
'Why are the little wheels turning?'

c. *Gdzie zgubił* <u>*się*</u> *bucik*? **M** (1;9)
where lose$_{\text{PAST.MASC.SG}}$ REFL shoe $_{\text{NOM.MASC.SG}}$
'Where did the shoe get lost?'

d. *Co zbiło* <u>*się*</u>? **W** (2;5)
what $_{\text{NOM.NEU}}$ break $_{\text{PAST.NEU}}$ REFL
'What broke?'

The classical GB analysis for (4) is NP-movement of the internal argument, leaving a trace: *The window opened t* (Burzio (1986) based on Perlmutter (1978)). More recent approaches include Reinhart (1996) and Hale and Keyser (1993 and later). Reinhart (based on Chierchia 1989) proposes a transitive entry as basic, as in (5a), with a lexical process of argument suppression yielding intransitive verbs that project only one argument as subject, as in (5b).

(5) a. *Basic verb entry* V ($\theta_1\theta_2$), e.g. open ($\theta_1\theta_2$)
b. *Suppression* R(V (x )), e.g. *R*(open (door))
The door opened.→ UNACCUSATIVE

Hale and Keyser take the intransitive entry as basic, with possible argument augmentation and transitivization (in the syntax). We remain agnostic about the adult grammar analysis, but the UEAPP renders Hale and Keyser's approach the most suitable for Polish child grammar. Our approach suggests that the early unique argument constructions with *się* are *augmented* as the grammar develops: the predicate projects (an) additional argument(s), establishing the internal-external argument distinction once the UEAPP is modified. The intransitive verbal entry can be analyzed as a more basic one, since it displays a one-to-one relation between verb and argument sister nodes.

Inherents are as in (6), with an obligatory clitic and possibly an oblique complement (*bać* <u>*się*</u> *wilków* 'fear$_{\text{INF}}$ REFL wolves$_{\text{GEN}}$'), and are often viewed as unaccusatives (nonagentive Nom NP). In Polish, they also combine with agentive NPs, thereby they can also be unergative, including *play* in (5c), *cough*, and *pray*.[1] The traditional analysis of unergatives is with Nom NP as external argument, and no movement.

(6) a. *Boję* <u>*się*</u>. **M** (1;7)
fear$_{\text{PRES.1SG}}$ REFL
'I am afraid. / I fear.'

b. *Pali* *się*. **B** (1;7)
be+on$_{PRES.3SG}$ REFL
'It is on. / It shines.'

c. *Bawimy* *się*. **B** (1;11)
play$_{PRES.1PL}$ REFL
'We play. / Let us play.'

A use with agentive NPs, common in Slavic, is the telic marker in (7).

(7) a. *Chcą* *się* *na* *-pić*? **W** (2;9)
want $_{PRES.3PL}$ REFL PPr -drink $_{INF}$
'Do they want to drink?'

b. *Nie na -uczyłem się pływać.*
Neg PPr -learn $_{PAST.MASC.1SG}$ REFL swim $_{INF}$
*bo za mały jestem.* **W** (2;9)
because too small $_{NOM.MASC}$ be $_{PRES.1SG}$
'I did not learn how to swim because I am too small.'

c. *Na- piję* *się*. **B** (1;7)
PPr drink $_{PRES.1SG}$ REFL
'I'll drink up.'

Clitic *się* combines with a P(erfective) Pr(efix) (*na*-) to indicate a delimited situation, and detransitivizes V.[2] The perfectivizing function is not frequent in our files, but can be found early ((7c) is Stage II), fitting the predicate modifier status of *się* in Stage II.[3] Thus far, then, the Polish clitic can combine with unaccusatives (nonagentive Nom) and unergatives (agentive Nom). This situation, then, a) challenges a common assumption in generative grammar, which is that reflexive clitic constructions always have unaccusative analyses (Marantz 1984 and later); b) suggests syntactic unification of unaccusatives and unergatives; and c) makes viable a predicate modifier view suitable for the UEAPP in early Polish.

The reflexivity function is exemplified in (8) and (9). For some, reflexive constructions are one-argument intransitives (Marantz 1984; Reinhart 1986 for different versions of this idea). The traditional view in GB (Chomsky 1981) is with two independent syntactic arguments.

(8) a. *Tata* *goli* *się*. **B** (1;7)
daddy$_{NOM}$ shave$_{PRES.3SG}$ REFL
'Daddy shaves.'

b. *Kąpią* *się*. **B** (1;7)
bathe$_{PRES.3PL}$ REFL
'They are bathing.'

(9) a. *Ja* *się* *wycieram.* K (2;4)
I$_{\text{NOM}}$ REFL wipe$_{\text{PRES.1SG}}$
'I am wiping myself.'

b. *Ja* *zamknąłem* *się* *na klucz* *w domu.* W (2;8)
I$_{\text{NOM}}$ lock$_{\text{PST.1SG.MASC}}$ REFL on key$_{\text{ACC}}$ in home$_{\text{PREP}}$
'I locked myself in the house.'

c. *Ja* *te* *żyletki* *powyciągałem* *i się*
I$_{\text{NOM}}$ these razors$_{\text{ACC}}$ take+out$_{\text{PST.1SG.MASC}}$ and REFL
*skaleczyłem.* K (2;4)
cut$_{\text{PST.1SG.MASC}}$
'I took these razors out and cut myself.'

On the first view, clitic constructions can be seen as predicative with modifier *się*, resembling unergatives and unaccusatives – not a new idea. The second view may receive support in Polish from the alternation of (8a) with *Tata goli siebie* 'Daddy is shaving himself', with a non-clitic Acc reflexive in A-position. Then, it can be assumed that the reflexive clitic can receive an analysis that is similar to the one often proposed for pronominal clitics. That is, the reflexive clitic heads the syntactic (clitic) chain of an argument independent from the Nom NP, and thus resembles SELF in the strategy Reuland (2001: 480) calls *protective* (i.e. argument positions are preserved). In Section 3, we combine the two views to capture the split chronology for the development of reflexivity in CHILDES.

Constructions in Stage III are impersonals, Dative impersonals, and reciprocals. Impersonals are as in (2) and (10), and differ in morphology from other patterns.

(10) a. *A* *tego* *wagonu* *się* *nie* *da.* K (2;4)
but this wagon$_{\text{GEN. MASC. SG}}$ REFL not be+able+to+do$_{\text{3SG}}$
'But one cannot {do/manipulate} this wagon.'

b. <**Tą*> *lokomotywę* *też* *się* *podłącza* K (2;4)
<*this$_{\text{INS}}$> locomotive$_{\text{ACC. FEM. SG}}$ also REFL connect$_{\text{PRES.3SG}}$
'One also connects this locomotive.'

First, impersonals cannot display an overt nominative NP, since the construction displays nominative properties that are attributed to the clitic itself (for Polish Kański 1986, for Italian Cinque 1988, among others). Second, V has default morphology: 3SG in the present tense, and Neuter in the past tense. In other uses, V agrees with the (overt) Nom NP. Third, with transitive Vs, impersonal patterns may display an Acc NP in affirmative clauses – (2) and (10b) – and a Gen NP in negative clauses – (10a).[4] The traditional view is that

Polish impersonals are active sentences (Dziwirek 1994 for references), and for Rivero (1999, 2000) they project two syntactic arguments. The {Acc/Gen NP} is the internal argument, and the clitic stands for the external argument, which is a type of SE-anaphor, as in Reinhart and Reuland (1993). More precisely, a (null) Nom pronoun with a human feature and no phi-features that does not make the predicate semantically reflexive raises from the VP to check Case against the clitic, repairing its referential deficiency. Crucial aspects for us are the two syntactic arguments, beyond the scope of the UEAPP, and the clitic as (argument) pronoun.

Patterns with clitics and Datives as in (11) are called involuntary state constructions in the academy grammar (Mieczkowska & Zondek 1984), inversions in relational grammar (Dziwirek 1994; Moore & Perlmutter 2000), and Dative impersonals here.

(11) a. *Pij, małpkusiu, pij, chce ci <u>się</u> pić?* **I** (2;9)
drink$_{\text{IMP.SG}}$ monkey$_{\text{VOC}}$ drink$_{\text{IMP.SG}}$ want$_{\text{PRES.3SG}}$ you$_{\text{DAT}}$ REFL drink$_{\text{INF}}$
'Drink, little monkey, drink, are you thirsty?'

b. *Nie uda <u>się</u> tatusiowi zrobić półki nowej.* **W** (2;8)
Neg be+able$_{\text{FUT.SG}}$ REFL daddy$_{\text{DAT}}$ make$_{\text{INF}}$ shelf$_{\text{GEN.SG.FEM}}$ new$_{\text{GEN.SG.FEM}}$
'Daddy will not {be able/manage} to make a new shelf.'

c. *Bo mi <u>się</u> już tego nie chce.* **W** (3;0)
because I$_{\text{DAT}}$ REFL already this$_{\text{GEN}}$ Neg want$_{\text{PRES.3SG}}$
'Because I do not want this any more.'

They are formally identical to (10) except for the Dative, with a) an obligatory (Nom) reflexive, b) a default morphology V, c) no overt Nom NP, and d) possibly an Acc or Gen object, as in (11b).[5] For them, we adopt the analysis in Rivero and Sheppard (2001). The Dative is equivalent to a Left Dislocated or high Applicative Phrase, and the semantic subject of predication. The remaining structure is a Tense Phrase that is the complement of the Applicative and is the predicate, and Nom *się* is the resumptive pronoun for the dislocated Dative. In this analysis, (11b) is similar to 'As to *Daddy*, *he* will not manage to make a new shelf', with two differences. The first is the Polish dislocated phrase in the Dative, and the second is the Polish resumptive, which is Nom *się*, not an ordinary pronoun. For us, the important aspects are the two syntactic arguments, and the clitic that functions as the pronoun that signals one of them.

In the reciprocal in (12), the clitic may alternate with non clitic *siebie*: *Żeby one siebie widziały* 'So that they see each other', which suggests two syntactic arguments similar to English *each other*. On this view, the Nom NP is the plural antecedent called distributor in Heim, Lasnik and May (1991), and the clitic signals an argument chain qua reciprocator.

(12) a. *Kogutki tu są, biją się dwa takie.* **I** (2;5)
rooster$_{\text{NOM.PL}}$ here are beat$_{\text{PRES.3PL}}$ REFL two$_{\text{NUM}}$ such$_{\text{DEM.PL}}$
'There are roosters here, they are fighting, these two.'

b. *Żeby one się widziały.* **W** (3;2)
so+that they$_{\text{NOM.PL.FEM}}$ REFL see$_{\text{PAST.3PL.NON-V}}$
'So that they (=two tape recorders) see each other.'

Finally, two uses absent in the database are middle and passive. The middle in (13) is rare and restricted to a few predicates.

(13) *Te samochody prowadzą się łatwo.*
These cars$_{\text{NOM}}$ drive$_{\text{PRES.3PL}}$ REFL easily$_{\text{ADV}}$
'These cars drive easily.'

Many consider the passive in (14) obsolete (Siewierska 1988, among others).[6] In most circumstances Polish adults opt for the impersonal or {Acc + Default V} pattern in (11), judging the passive or {Nom + agreeing V} pattern in (14) as unnatural or ungrammatical.

(14) *Dom szybko się zbudował.*
house$_{\text{NOM.SG.MASC}}$ fast$_{\text{ADV}}$ REFL built$_{\text{PAST.SG.MASC}}$
'The house was built fast.'

With asymptomatic morphology, the choice most often reflected in CHILDES is the impersonal, as in (15), translated by generic **you**: 'You join it like this'.

(15) *Tak się podłącza.* **K** (2;4)
so REFL connect$_{\text{PRES.3SG}}$
'ONE connects (it) like this.' NOT: '(It) is connected like this.'

In sum, the Polish clitic is a functional category poor in semantic content, and unusually rich in syntactic functions. For acquisition purposes, however, these multiple functions can be divided into two general classes: a) non-argument predicate modifier, and b) (defective) argument. Telic and Dative impersonal uses have received limited attention in generative grammar, but fit into the two classes. Regarding acquisition, non-argument functions emerge first and argument functions follow, due to the UEAPP, which undergoes maturation.

The UEAPP also guides the development of modal *wolno*, which is richer in semantic content but more limited in syntactic functions.

## 3. Stages in the acquisition of Polish reflexive clitics

Children are systematically exposed to the major clitic constructions, but some adult patterns emerge soon, and others are systematically delayed, independent of input.

Figure 1 sketches the three grammatical Stages of **I** and **W**, the only children in our corpus who display all three stages. The measure of grammatical development is the mean length of utterance (MLU), which is said to be a reliable method of identifying the point of development of grammar in children (Brown 1973). Brown carried MLU calculations, recording the mean number of morphemes per utterance when analyzing his children's linguistic development. He hypothesized that, if done consistently, MLUs of children at a similar stage of grammatical development will be similar. Using CLAN software provided with the CHILDES database, we carried out calculations of mean length of utterance *in words*. Polish is a highly inflected language. Comparing its morphemic calculations with those of children acquiring English, it contains considerably more morphemes per word, an observation that made us turn to calculations of the number of words, not morphemes, per utterance.

As Figure 1 suggests, we did not find a correlation between **I**'s and **W**'s MLU values. The children and their respective adult input values did, nevertheless, correlate. The child whose caretakers' utterances contained more words also spoke more. Interestingly, the trend in the children's MLU values is ascending in relation to the more stable values in the corresponding adult input. This indicates that sentences get longer, and grammar develops independent of the length of input utterances. Note also that the age disparity in the stages of reflexive development in **I** and **W** diminishes considerably by Stage III.

Wawrzon's rocky road of acquisition of the *się*-constructions is illustrated in Figure 2. The emergence of the early reflexive constructions is marked by a sharp peak at Stage II, after which the percentage of correct *się* use rises, reaching a plateau at Stage III.

Let us look at the stages of development in detail. During the first months of development, in Stage I, the required clitic is missing in unaccusatives including inchoatives (16–17), inherents of unaccusative and unergative types, (18–19), telics (20), reflexives (21), and impersonals (22).

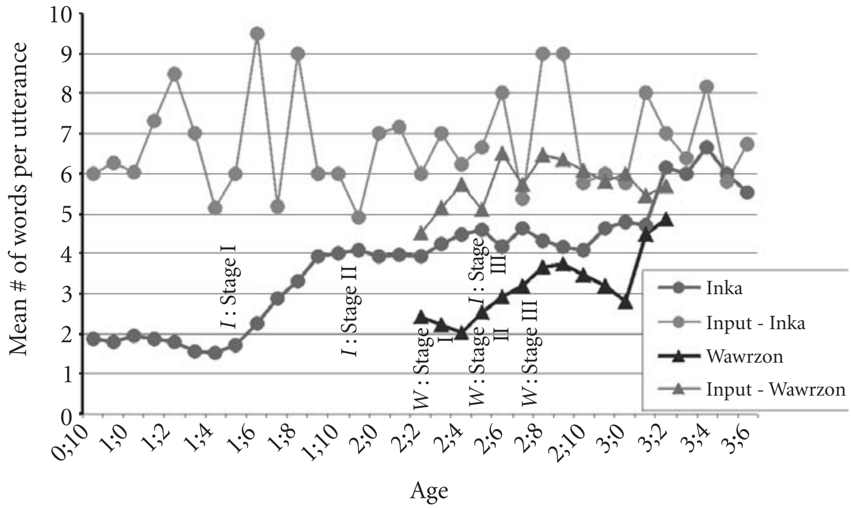



**Figure 1.** The Mean Length of Utterance (MLU) in words for Inka (I) and Wawrzon (W) and their input – all stages

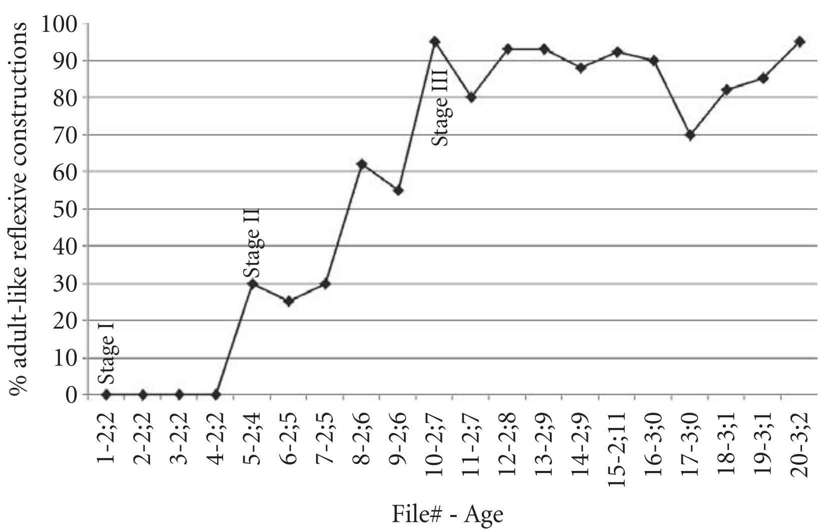



**Figure 2.** Wawrzon's overall success with reflexive constructions

(16) **Misiu huśta.* **I** (1;7)
bear$_{\text{NOM}}$ swing$_{\text{PRES.3SG}}$
'The teddy bear is swinging.' *Adult*: Misiu huśta *się*.

(17) **Złamało.* **W** (2;2)
break$_{\text{PAST.SG.NEU}}$
'It broke.' *Adult*: Złamało *się*.

(18) a. **Boi śnieg.*[7] **I** (1;8)
fear$_{\text{PRES.3SG}}$ snow$_{\text{NOM.SG.MASC}}$
'He is afraid of snow.' *Adult*: Boi *się* śniegu$_{\text{GEN}}$.
b. **Przestraszył też piesek.* **W** (2;2)
scare$_{\text{PAST.SG.MASC}}$ also doggy$_{\text{NOM.SG.MASC}}$
'The doggy got scared also. *Adult:* Przestraszył *się* też piesek.

(19) **Bawił będę.* **W** (2;2)
play$_{\text{PPART.SG.MASC}}$ be$_{\text{FUT.1SG}}$
'I will play.' *Adult*: Bawił *się* będę.

(20) **Na-pić mogę?* **W** (2;2)
PP-drink$_{\text{INF}}$ can$_{\text{PRES.1SG}}$
'Can I have a drink?' *Adult*: Na-pić *się* mogę?

(21) a. **Dziaduś goli.* **I** (1;8)
grandpa$_{\text{NOM}}$ shave$_{\text{PRES.3SG}}$
'Grandpa shaves.' *Adult:* Dziaduś *się* goli.
b. **Nie kłuję.* **W** (2;2)
Neg prick$_{\text{PRES.1SG}}$
'I will not prick myself.' *Adult*: Nie ukłuję *się*.

(22) **Tak składa.* **W** (2;2)
such fold$_{\text{PRES.3SG}}$
'One folds it like this.' *Adult:* Tak *się* składa.

Between the ages of 1;10 and 2;4 the clitic emerges, marking Stage II of the acquisition. Specifically, the children exhibit Stage II grammatical patterns in the following age intervals: **I**: 1;10 – 2;4; **W**: 2;4 – 2;5; **K**: 2;1 – 2;3; **B**: 1;7 – 1;11; and **M**: 1;7 – 1;10.

In this stage, the clitic emerges with unaccusatives as in (1), *Przewróciłem się* 'I fell', *Okno się otworzyło* 'The window opened', unaccusative and unergative inherents as in (6), *Boję się* 'I fear' and *Bawimy się* 'We play', and in the telic function in (7c) *Na-piję się* 'I'll drink up'. Omissions are still found, but interesting novelties are the overgeneralizations or "wrong" uses that emerge for each of these early functions. The third interesting feature is the reflexive use that emerges only in some children and with a specific type of predicate.

Let us examine these features in turn. We begin by illustrating in (23) Stage II omissions.

(23) Unaccusative

a. **Proszę, tatuś, obudź.* **M** (1;9)
please$_{\text{PRES.1SG}}$ daddy$_{\text{NOM}}$ wake$_{\text{IMP.SG}}$
'Please, daddy, wake up.' *Adult*: Proszę, tatuś, obudź *się*.

b. **Chowam.* **B** (1;7)
hide$_{\text{PRES.1SG}}$
'I am hiding.' *Adult:* Chowam *się*.

c. **Przewrócił.* **K** (2;2)
fall$_{\text{PST.3SG.MASC}}$
'(He) fell.' *Adult*: Przewrócił *się.*

Reflexive – intrinsic

d. **Goli.* **B** (1;7)
shave$_{\text{PRES.3SG}}$
'(He) is shaving.' *Adult:* Goli *się.*

Reflexive – extrinsic

e. **Uderzyłem.* **B** (1;7)
hit$_{\text{PST.1SG.MASC}}$
'I hit myself.' *Adult*: Uderzyłem *się*.

Inherent

f. **Bawiłem.* **K** (2;3)
play$_{\text{PST.1SG.MASC}}$
'I was playing.' *Adult*: Bawiłem *się*.

g. **Pali.* **B** (1;7)
be+on$_{\text{PRES.3SG}}$
'It is on.' *Adult*: Pali *się.*

We attribute to the UEAPP the early appearance of *się* with unaccusatives, the two types of inherents, and the {telic/excess} use. Under the UEAPP, (a) each predicate establishes a biunique relation with an argument regardless of type; (b) the clitic behaves like an adverb that adds a specification without disturbing this relation; and (c) unaccusatives and unergatives are similar in terms of the predicate-argument relation. If the UEAPP is adopted, the best analysis of the predicate for this period is with a basic intransitive shell (as in Hale & Keyser 1983 and later), and the transitive shell with argument suppression looks less desirable. Recall Reinhart's (1996) proposal of argument suppression illustrated in (5) in Section 2: the transitive verbal entry undergoes suppression to yield an unaccusative verb. We believe the intransitive entry serves as a basis for argument augmentation.

The proposed analysis can also account for three types of overgeneralizations that arise in Stage II, which continue in Stage III. They all fit the UEAPP, and the hypothesis that unaccusatives and unergatives belong to the same syntactic type, and are of three types. The first is with unaccusatives, (24), and the second with unergatives, (25).

(24) a. **Książeczka spada się.* **B** (1;8)
book NOM.DIM.SG fall$_{\text{PRES.3SG}}$ REFL
'The little book is falling.' *Adult:* Książeczka spada.

b. **Na nogi mi się wykipiało.* **W** (2;8)
On legs me$_{\text{DAT}}$ REFL boil+over$_{\text{PST.3SG.NEU}}$
'It boiled over on my legs.' *Adult*: Na nogi mi wykipiało.

(25) a. **Pływała się.* **B** (1;8)
swim$_{\text{PST.3SG.FEM}}$ REFL
'(She) swam.' *Adult*: Pływała.

b. **Ty się nie wychodź!* **W** (2;7)
You REFL Neg come+out$_{\text{PRES.3SG}}$
'You do not come out!' *Adult:* Ty nie wychodź.

c. **Ja tu się wejdę do komody.* **W** (2;7)
I$_{\text{NOM}}$ here REFL go+in$_{\text{FUT.1SG}}$ to cupboard$_{\text{GEN}}$
'I'll go into the cupboard.' *Adult*: Ja tu wejdę do komody.

The third is with the {telic/excess} use(s). *Lala* in (26c) is an onomatopeic invention, and (26e), an unaccusative, has the wrong tense.[8]

(26) a. **Na- pisze się.* **B** (1;7)
PPr- write$_{\text{PRES.3SG}}$. REFL
Intended: 'He (= a man writing) is writing.' *Adult*: Pisze.

b. **Pisze się?* **B** (1;7)
write$_{\text{PRES.3SG}}$ REFL *Adult*: Pisze?
'Is it (= the tape recorder) {writing, recording}?'

c. **Się lala.* **B** (1;7)
Intended: 'It (=the tape recorder) is making noise.'

d. **Widzisz się.* **K** (2;3)
see$_{\text{PRES.2SG}}$ REFL
'You see.' *Adult*:Widzisz.

(cf. the grammatical intensifying use of the reflexive in *Patrzysz się* 'You look.')

e. **Ja się tak, o taki, o tak byłe taki*
I$_{\text{NOM}}$ REFL so oh like+this, oh so be$_{\text{PAST.1SG.MASC}}$ so
*wielki.* **K** (2;6)
big
Intended: 'I will grow to be this big.'

Turning to the reflexive use, it is surprisingly infrequent in both children and adults, so it must emerge under limited exposure. Some percentages are as follows. **B** hears 146 clitics with 5 (3.4%) in the reflexive function, and 2 repeat what he said. **M** hears 202 clitics, with 6 (3.0%) for reflexivity. By contrast, impersonals are more frequent but not produced: **B** hears 29 (close to 20.0%), and **M** hears 23 (11.4%). At 1;7, **B** utters *Nie da **się*** 'It cannot be done / One cannot do this' twice after his father, which we count as prompted expression, not as an impersonal type (10a).

**B**, **I**, and **M** develop the reflexive use in Stage II with what are called intrinsically reflexive predicates in the literature, as in (8a) *Tata goli się* 'Daddy shaves', and (27). The class awaits definition, but parallels the use of the simple reflexive *zich* as SE-anaphor in Dutch (Everaert 1986), *sig* in Icelandic and Danish, and contains verbs that may lack *SELF* in English. Its early appearance with the clitic suggests an intransitive analysis for reflexivity in Stage II, complying with the UEAPP.

(27) a. *Ubiera się.* **M** (1;9)
dress$_{\text{PRES.3SG}}$ REFL
'(She) is dressing.'

b. *Pimpek się myje.* **I** (1;11)
Pimpek REFL wash$_{\text{PRES.3SG}}$
'Pimpek is washing.'

In Stage III, **K** and **W** develop the reflexive use with predicates that need not be intrinsically reflexive in English or Germanic, which are also found in **I** at this time, as in (9a) *Ja **się** wycieram* 'I am wiping myself', and (28).

(28) a. *Tak jak Witek się poo-bijał.* **W** (2;7)
just like Witek$_{\text{NOM}}$ REFL Pf-hit$_{\text{PST.3SG.MASC}}$
'Just like Witek hit himself (all over).'

b. *Niech się Tenia pokaleczy, ja nic nie mówię,*
let REFL Tenia cut$_{\text{PRES.3SG}}$ I$_{\text{NOM}}$ nothing$_{\text{ACC}}$ Neg say$_{\text{PRES.1SG}}$
*niech się pokaleczy.* **I** (3;2)
let REFL cut$_{\text{PRES.3SG}}$
'Let Tenia cut herself, I'm not saying anything, let her cut herself.'

The delay of the reflexive function to Stage III in some children suggests a transitive analysis as alternative. On this view, *się* plays the equivalent of what Reuland (2001) labels a "protective function," and indicates an argument position in syntax. In contrast with the intrinsic, non-argument reflexive analysis proposed for Stage II, *się* in reflexive constructions in Stage III denotes an internal argument projected in the syntax.

If in Stage II reflexivity develops via an intransitive analysis, a fourth type of wrong use of the clitic may receive an interesting account. During this period, **B** experiences difficulty expressing clitics with body parts as in (29), which other children usually do not mention:

(29) a. **Tu* *uderzyłem* *<u>się</u>* *w...* **B** (1;7)
here hit$_{\text{PST.1SG.MASC}}$ REFL on... = *Here, I hit myself on...
Intended in context: Uderzyłem *się* w nogę. 'I hit myself on the leg.'

b. **Wykręciłem* *<u>się.</u>* **B** (1;11)
twist$_{\text{PST.1SG.MASC}}$ REFL = *I twisted myself.
Intended in context: Wykręciłem sobie$_{\text{DAT}}$ nogę$_{\text{ACC}}$. 'I twisted my leg.'

If body-part reflexives involve two syntactic arguments, they should cause problems in Stage II, and mature in the proper way only in Stage III. Examples in (29) illustrate the problems encountered at this stage: (29a) is an incomplete sentence where only the preposition but not its complement is expressed; (29b) lacks the body part affected by the action of twisting.

In sum, the reflexive function may develop in Stage II via non-argument *się*, or in Stage III together with reciprocals, impersonals and Dative impersonals, which involve argument *się*. As mentioned above, we propose that the reason for late emergence of extrinsic reflexives is the construction's incompatibility with the UEAPP. Recall that the proposed Stage I contains no reflexives of any types. Then, Stage II is marked by the application of the UEAPP, and sees constructions involving a simple, biunique relation between the verb and its argument. It is then that unaccusative (4), inherent (6), and intrinsic reflexive (8) constructions emerge. Finally, upon reformulation of the UEAPP at Stage III, the grammar is able to accommodate two arguments per predicate. Consequently, the children utter reflexives of the extrinsic reflexive (9), Dative (11), reciprocal (12), and impersonal (15) types.

To conclude, early clitic constructions fall under the UEAPP and late clitic constructions do not.

## 4. The development of *wolno*

In this section, we examine the development of *wolno*, based on the files of **I** and her caretakers. We examined the files of all children, but others offer limited data. **B** utters a total of two *Nie wolno*. **K** has three without Dative, including *Nie wolno się bawić radiem* 'It is not allowed to play with the radio' at 2;3. **M** lacks *wolno*. **W** has one without Dative at 3;1. By contrast, **I** has 75 *wolno*-sentences from 0;10 to 3;6, the last file examined in our study.

As mentioned in the introduction, in adult grammar the full-fledged argument structure of *wolno* is with two arguments: topic of permission and Dative goal, as in (3a) [*Mnie*] *nie wolno* [*tak daleko iść*] *prawda?* 'I am not allowed to walk that far, right?'. Inka is systematically exposed to this construction from Stage I, but her production is delayed until Stage III. In our view, this modal predicate falls under the UEAPP, so the two arguments can only be expressed in Stage III, when this early principle is modified.

*Wolno* emerges in **I** at 1;9, coinciding with the clitic, which was dubbed Stage II. **I**'s first utterance with the modal is (3b) *Tatusiu, nie wolno bić mamusi* 'Daddy, it is not allowed to hit mommy', and shows the symptoms of Stage II, which lasts until 3;0 for the modal. In Stage II, the goal is expressed as a Vocative, not as a Dative argument of the predicate. Thus, in this period, *wolno* displays at most one syntactic argument in tune with the UEAPP. Recall that the late emergence of the second argument with *wolno* is not due to the lack of A-chains: we believe that the predicate's external argument requirement is the crucial factor in the delay. The developmental delay with the second argument of *wolno* arises as a consequence of the UEAPP. The proto-principle must be reformulated to accommodate the Dative argument projected by the modal predicate. Here are some other examples.

(30) a. *Nie wolno brzydko mówić, ty Barbariba!* **I** (2;1)
Neg allow ugly$_{\mathrm{ADV}}$ talk$_{\mathrm{INF}}$, you$_{\mathrm{NOM}}$ Barbariba$_{\mathrm{NOM}}$
'To talk ugly is not allowed,you Barbariba!'

b. *Ty Jola, nie wolno na ulicę chodzić!* **I** (2;4)
you Jola, Neg allow on street walk$_{\mathrm{INF}}$
'You Jola, to walk on the street is not allowed!'

c. *Nie wolno misiu do pieluszki sisiać!* **I** (2;6)
Neg allow bear$_{\mathrm{VOC}}$ to diaper$_{\mathrm{GEN}}$ pee$_{\mathrm{INF}}$
'Bear, to pee in the diaper is not allowed!'

Another interesting feature is that if the required argument is nonovert, as in (31), often the previous discourse provides an antecedent. This suggests that the argument is syntactically present as a null pronoun whose content is recovered in context.

(31) a. *Nie szczekaj, nie* <u>*wolno!*</u> **I** (2;1)
Neg bark$_{IMP.2SG}$ Neg allow
'Do not bark! It is not allowed.'

b. *Mietek, czemu śpisz moje łóżko, nie* <u>*wolno*</u> **I** (2;1)
Mietek, why sleep$_{PRES.2SG}$ my bed, Neg allow
'Mietek, why are you sleeping on my bed? It is not allowed.'

c. *Nie sprzedaj Inki tatuniu, nie* <u>*wolno*</u> **I** (2;2)
Neg sell$_{IMP.2SG}$ Inka$_{GEN}$, daddy$_{VOC}$, Neg allow
'Do not sell Inka, Daddy! It is not allowed.'

Datives are missing in *wolno* constructions in Stage II, but they can be found during this period as in (32), so the Dative per se is not the source of the delay:

(32) *Nie* <u>*wolno*</u>*, żaba, mówiłam ci.* **I** (2;2)
Neg allow frog$_{NOM}$ tell$_{PST.1SG.FEM}$ you$_{DAT}$
'It is not allowed, frog, I told you!'

In this stage, Inka comes to differ in style from the adults surrounding her, as the following numbers demonstrate.

Stage I lasts from 0;10 to 1;8 when the child utters no *wolno,* and hears 94 tokens: 12 (12.7%) with a Vocative as in (30), and 10 (10.6%) with a Dative as in (3a). Stage II is from 1;9 to 2;9, when she utters 55 tokens, 16 (29.0%) with a Vocative, more than doubling the adult rate for Stage I, and 0 with a Dative. By contrast, in this period adults utter 39 tokens, no Vocatives, and 22 Datives (56.0%!!). Why should the behavior of adults change from Stage I to Stage II and why should Vocatives be suppressed? Why should Inka be ignorant of the

**Table 3.** The emergence of <u>wolno</u> in Inka vs. adult input

| Stage; Age | Inka: wolno | | | Input: wolno | | |
|---|---|---|---|---|---|---|
| | Total tokens | Tokens with vocative (% total) | Tokens with dative (% total) | Total tokens | Tokens with vocative (% total) | Tokens with dative (% total) |
| Stage I; 0;10–1;8 | 0 | – | – | 94 | 12 (12.7) | 10 (10.6) |
| Stage II; 1;9–2;9 | 55 | 16 (29.0) | 0 | 39 | 0 | 22 (56.0) |
| Stage III; 3;0–3;6 | 29 | 1 (3.4) | 8 (27.6) | 22 | 1 (4.5) | 16 (72.7) |

Dative input in Stage II? We submit that on the one hand, Vocatives make the child sound rude, as glosses convey, so adults try to influence the child and suppress them, favoring the more refined Dative. On the other hand, the child is under the influence of the UEAPP so she ignores the Dative input.

Stage III goes from 3;0 to 3;6, which is the last file we examined, and Inka produces 29 tokens: 1 with a correct Vocative (3.4%), and 8 with Datives (27.6%). Thus, from Stage II to Stage III she reverses the proportion of Datives and Vocatives, which we take as a signal of the modification of the UEAPP. However, the *wolno*-constructions she hears in Stage III follow the proportions of those heard in Stage II: out of the 22 heard, 16 constructions (72.7%) are with a Dative. Here are some late examples:

(33) a. *A* <u>*wolno*</u> [*mi*] [*się* *bawić*] *mamusiu?* **I** (3;0)
but allow I$_{DAT}$ REFL play$_{INF}$ mommy$_{VOC}$
'But, am I allowed to play, mommy?'

b. <u>*Wolno*</u> [*mi*] [*się* *bawić* *z* *dziewczynkami*]*?* **I** (3;0)
allow I$_{DAT}$ REFL play$_{INF}$ with girls$_{INST}$
'I am allowed to play with girls?'

c. *A* *nie* <u>*wolno*</u> [*jej*] [*krzyczeć*], *będzie* *płakała.* **I** (3;4)
but Neg allow she$_{DAT}$ yell$_{INF}$ will cry$_{P.PART.SG.FEM}$
'But she is not allowed to yell, she will cry.'

In brief, the development of the modal *wolno* is similar to that of the clitic *się*, strengthening the evidence for the existence of UEAPP in child grammar. While the proto-principle is operative, *wolno* projects only one argument in the syntax. A reformulation of the proto-principle allows the second argument to surface.

## 5. Conclusion

We have argued for Maturation in view of the UEAPP that establishes a biunique relation between predicates and arguments. In Polish, the development of reflexive clitics and modal *wolno* 'allow' observed in five children in the CHILDES database supports Maturation: UG is not available in adult form in all stages of acquisition, and matures. Reflexive clitic and modal constructions develop in three chronologically parallel stages. The first is complete omission or absence. The second obeys the U(nique) E(xternal) A(rgument) P(roto)-P(rinciple). Unaccusative, unergative, and intrinsic reflexive uses of the clitic develop, and *wolno*-constructions display only one syntactic argument. The proto-principle is modified in the third stage, when clitic and modal constructions with two syntactic arguments emerge: the children utter *się* -impersonals, reciprocals, Datives, and extrinsic reflexives, as well as *wolno*-Datives. In brief, the UEAPP regulates the syntactic development of functional categories such as the reflexive clitic *się* and modal *wolno*, supporting the Maturation Hypothesis.

## Notes

* Research for this paper was partially supported by SSHRCC Research Grant 410-2000-0120 to María Luisa Rivero. Preliminary versions were read at Psycholinguistics Shorts, University of Ottawa, in November 2000, and the Asymmetry Conference, University of Quebec in Montreal, in May 2001. We are grateful to the audiences for many useful questions and comments. Special thanks go to Helen Goodluck and Danijela Stojanović for much help and advice as this research progressed.

**1.** Similar French Vs do not combine with *être* 'be', and Spanish Vs are not used as Absolute Participles, proposed as unaccusativity symptoms: *L' enfant a (*est) joué* 'The child (has/*is) played,' **El niño jugado, entraron* '*The child played, (they) went in.'

**2.** Spanish has also a telic *se* which contrasts in (in)transitivization, coexisting with an obligatory object NP as delimiter (Nishida 1994; Sanz 1999):

i. a. *Se aprendió la lección.*
REFL learn$_{\text{PAST.3SG}}$ the lesson
'He learned the lesson (well, completely).'

b. *Se bebió el vino.*
REFL drink$_{\text{PST.3SG}}$ the wine
'He drank up (all) the wine.'

**3.** Another (agentive) predicate modifier use is in (i).

i. a. *Bo ja pytam się.* **W** (2;6)
because I$_{\text{NOM}}$ ask$_{\text{PRES.1SG}}$ REFL
'Because I am asking.'

b. *Babunia też się na tym nie zna.*
granny$_{\text{NOM}}$ also REFL about this Neg know$_{\text{PRES.1SG}}$
'Granny has no idea about it.' (from **M**'s grandmother)

Vs like *ask* and *know* display non-reflexive and reflexive versions with similar (truth conditional) semantics, with the last indicating deeper involvement. Grappin (1963) glosses *patrzeć* as 'to look' and *patrzeć się* 'to look intensely'. Reflexive *look*, which is compatible with Stage II, develops in **I** at 1;6, but it is unusually frequent in her input, so we consider it prompted. The productive and non-prompted use of *się* begins in Inka at 1;10.

**4.** In (b), the Demonstrative has the wrong Instrumental.
The transitive impersonals in the text seem to develop slightly earlier than intransitive impersonals of type (i).

i. *Tu* *się* *wsiada* *i* *jedzie* *się.* **W** (2;9)
here REFL get+on$_{PRES.3SG}$ and go$_{PRES.3SG}$ REFL
'Here one gets on and goes (by vehicle).'

If (i) contains only one syntactic argument indicated by the clitic, its delay does not follow from the UEAPP. Rivero (1999, 2000) argues that impersonals contain SE-anaphor chains with mixed A and A-bar properties, which could locate their delay in the ACDH discussed in the introduction. In the examined period some impersonal types have not developed. The unaccusative type for *One dies only once* is missing, but it is found in **I**'s files with *die* at 3;8, and with *grow* at 4;2.

**5.** The constructions in (11) are often grouped together with (inherently reflexive) Nom + Dative psychological Vs such as *please/like* as in (i).

i. *Ta* *mi* *się* *podoba,* *ta,* *ta.* **W** (2;7)
this$_{NOM.FEM.SG}$ I$_{DAT}$ REFL please tsubPRES.3SG this$_{NOM.FEM.SG}$ this$_{NOM.FEM.SG}$
'This one pleases me, this one, this one! / I like this one, this one, this one!'

The two differ in morphology, so for this and other reasons should not be unified. (i) relies on the modifier use, not on (impersonal) argument *się*, and falls outside of the UEAPP with Theme and Experiencer as syntactic arguments. We do not examine Nom-Dat psychological Vs, which are not always (inherently) reflexive, and examples with *please/like* are few. However, the UEAPP also seems to delay psychological Vs. Adult-like sentences as in (i) seem to belong to Stage III, and the early dativeless strategy in (ii) is reminiscent of *wolno* constructions in Stage II, with only one argument projected in the syntax:

ii. a. **Podoba* *się.* **M** (1;8)
please$_{PRES.3SG}$ REFL *Adult*: Podoba mi$_{DAT}$ *się.*
'It-pleases (me). / (I) like-it.'

b. **Podobają.* **I** (2;5)
please$_{PRES.3PL}$ *Adult*: Podobają mi$_{DAT}$ *się.*
'They-please (me). / (I) like-them.'

c. **Nie podoba.* **W** (2;4)
Neg please$_{PRES.3SG}$ *Adult:* Nie podoba mi$_{DAT}$ *się.*
'It does not please (me). / (I) do not like it.'

**6.** Passive reflexive constructions are common in other Slavic languages, and are often analyzed with NP-movement resulting in an A-chain. From this perspective, the comprehension experiments carried out by Stojanović (2000) with Croatian and Serbian *se*-passives are important for our purposes.
These experiments show that Croatian children (mean age 3;9) and Serbian children (mean age 4;2) understand *se*-passives such as (i) along parallel lines, but differ considerably from the adults tested as controls, which in our view is consistent with Maturation:

i. *Tamo* *se* *jedu* *jabuke.*
there REFL eat$_{PRES.3PL}$ apples$_{NOM}$
'There apples are eaten. / There people eat apples.'

When presented with the construction in (i) and three pictures in a sentence-picture matching task, 100% of the adults selected a picture where a boy is looking at a group of people eating apples. By contrast, 21% of the children chose a picture where the boy was looking at the apples, 39% chose a picture where the boy was eating an apple, and only 40% chose the picture unanimously selected by the adults.
The contrast between children and adults seems clear, and we suggest that the reason is the A-Chain Deficit Hypothesis. If children lack A-chains or experience difficulties with them, then (i) is probably treated as an adjectival passive, which makes it resemble in child grammar the unaccusative and unergative reflexive constructions of Polish. That is, the NP *apples* is in a predicative relation with *eat*, so the sentence corresponds to something along the lines of *These apples are fit for eating*, which can suit all three pictures, and accounts for the three-way split in the children. By contrast, for the adults the sentence involves NP-movement and an implicit Agent, which is the "passive" reading they all select.

**7.** In (18a), NP is Nom, not the expected Gen. This could be because the child projects it as the unique argument, similar to *Snow frightens*.

**8.** While the colloquial register allows for some variation in clitic position, usually Polish clitics are non-initial, a rule which children may ignore. The ungrammatical (i) could be interpreted as an attempt to repair the violation of the strict initial position prohibition.

i. **<u>Się</u>* *nie* *pali* *<u>się</u>*. **B** (1;7)
REFL Neg be+on$_{\text{PRES.3SG}}$ REFL
'It is not on.' (said of the tape recorder light). *Adult*: Nie pali *się*.

## References


Babyonyshev, M., Ganger, J., Pesetsky, D., & Wexler, K. (2001). The maturation of grammatical principles: evidence from Russian unaccusatives. *Linguistic Inquiry, 32*, 1–43.

Borer, H. & Wexler, K. (1987). The maturation of syntax. In T. Roeper and E. Williams (Eds.), *Parameter Setting* (pp. 123–172). Dodrecht: Reidel.

Borer, H. & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory, 10*, 147–189.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge: Harvard University Press.

Burzio, L. (1986). *Italian Syntax*. Dordrecht: Reidel.

Chierchia, G. (1989). A semantics for unaccusatives and its syntactic consequences. Ms., Cornell University.

Chomsky, N. (1964). *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Cinque, G. (1988). On *si* constructions and the theory of *arb*. *Linguistic Inquiry, 19*, 521–583.

Dziwirek, K. (1994). *Polish Subjects*. New York: Garland.

Everaert, M. (1986). *The Syntax of Reflexivization*. Dordrecht: Foris.

Grappin, H. (1963). *Grammaire de la langue polonaise*. Paris: Institut d'études slaves.

Hale, K. & Keyser, S. J. (1993). On argument structure and the lexical expression of syntactic relations. In K. Hale, and S. J. Keyser (Eds.), *The View from Building 20* (pp. 53–109). Cambridge, MA: MIT Press.

Heim, I., Lasnik, H. & May, R. (1991). Reciprocity and plurality. *Linguistic Inquiry, 22*, 63–101.

Hyams, N. (1987). The theory of parameters and syntactic development. In T. Roeper and E. Williams (Eds.), *Parameter Setting* (pp. 1–22). Dordrecht: Reidel.

Kański, Z. (1986). *Arbitrary Reference and Reflexivity: A Generative Study of the Polish Pronoun się and its English Equivalents*. Uniwersytet Śląski, Katowice.

Kubiński, W. (1987). *Reflexivization in English and Polish*. Tübingen: Niemeyer.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. 3rd Ed. Vol. 1 and 2. Mahwah, NJ: Lawrence Erlbaum.

Marantz, A. (1984). *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

Mieczkowska, Z. & Zondek, J. (Eds.) (1984). *Gramatyka współczesnego języka polskiego: morfologia*. Warsaw: Państwowe Wydawnictwo Naukowe.

Moore, J. & Perlmutter, D. M. (2000). What does it take to be a dative subject? *Natural Language and Linguistic Theory, 18*, 373–416.

Nishida, C. (1994). The Spanish reflexive clitic *se* as an aspectual class marker. *Linguistics, 32*, 425–458.

Perlmutter, D. (1978). Impersonal passives and the unaccusative hypothesis. *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*. 157–189.

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge: Harvard University Press.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

Reinhart, T. (1996). Syntactic effects of lexical operations: Reflexives and unaccusatives. *UIL OTS Working Paper*, Utrecht University.

Reinhart, T. & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry, 24*, 657–720.

Reuland, E. (2001). Primitives of binding. *Linguistic Inquiry, 32*, 439–492.

Rivero, M. L. (1999). On impersonal *się* in Polish: A simplex expression anaphor. *Journal of Slavic Linguistics, 8.1*, P. Banski and E. William (Eds.) [2000 = Volume on Polish Syntax].

Rivero, M. L. (2000). On impersonal reflexives in Romance and Slavic and semantic variation. In J.Camps and C. Wiltshire (Eds.), *Romance Syntax, Semantics and L2 Acquisition*. Amsterdam: John Benjamins.

Rivero, M. L. & Sheppard, M. (2001). On indefinite reflexive clitics in Slavic: Polish and Slovenian. To appear in *Natural Language and Linguistic Theory*.

Sanz, M. (1999). Aktionsart and transivity phrases. In E. Treviño and J. Lema (Eds.), *Semantic Issues in Romance Syntax* (pp. 247–261). Amsterdam: John Benjamins.

Siewierska, Anna (1988). The Passive in Slavic. In M. Shibatani (Ed.), *Passive and Voice* (pp. 243–289). Amsterdam: John Benjamins.

Stojanović, D. (2000). The clitic *se* and its numerous uses: Evidence from child grammar. Ms., University of Ottawa. Paper read at the *Workshop on Theoretical Problems of Pronouns*, Esse-2000, Helsinki.

Weist, R., Wysocka, H., Witkowska-Stadnik, K., Buczowska, R. H. E., & Konieczna, E. (1984). The defective tense hypothesis: on the emergence of tense and aspect in child Polish. *Journal of Child Language, 11*, 347–374.

Weist, R. & Witkowska-Stadnik, K. (1986). Basic Relations in Child Language and the Word Order Myth. *International Journal of Psychology, 21*, 363–381.

# A subgrammar approach to language acquisition

David Lebeaux
NEC Research Institute

In this paper, I introduce a new approach to language acquisition, called a subgrammar approach. This approach is an alternative to the parameter-setting approach[1]. Consider the pattern of growth in a grammar. At a particular time, conjunction "comes in". We may model this as a generalized transformation fusing two structures into one (Lebeaux 1988, 2000a, 2000b).

(1) conjunction:

s1: John
s2: Bill
→ John and Bill

s1: saw her
s2: talked to her
→ saw her and talked to her

At some other point, relative clause adjunction "comes in".

(2) Relative clause adjunction

s1: the man left
s2: who I know
→ the man who I know left

s1: I met the women
s2: who was tall
→ I met the woman who was tall

The acquisition sequence which underlies these changes for particular children looks like this (Miller 1973):

(3) Development of Conjunction (number used)

| Age | 2.2 | 2.3 | 2.5 | 2.7 | 2.9 | 3.1 | 3.6 |
|---|---|---|---|---|---|---|---|
| Susan | 1 | 0 | 9 | 16 | 19 | 11 | 43 |

| Age | 2.5 | 2.6 | 2.8 | 2.9 | 3.1 | 3.4 |
|---|---|---|---|---|---|---|
| Carl | 4 | 0 | 5 | 4 | 11 | 23 |

| Age | 2.4 | 2.5 | 2.6 | 2.7 | 2.8 | 2.10 | 3.0 | 3.1 | 3.6 |
|---|---|---|---|---|---|---|---|---|---|
| Christy | 2 | 1 | 0 | 3 | 0 | 3 | 2 | 6 | 16 |

As can be seen, conjunction "comes in" for Susan at about age 2.7, Carl controls it at 3.4, and Christy at 3.6.

The table for relative clause adjunction is the following:

(4) Relative Clause Adjunction (Susan)

| Age | 2.2 | 2.3 | 2.5 | 2.7 | 2.9 | 3.1 | 3.6 |
|---|---|---|---|---|---|---|---|
| what | 4 | 4 | 1 | 1 | 4 | 7 | 14 |
| who | 0 | 0 | 0 | 0 | 0 | 1 | 3 |
| that | 0 | 1 | 0 | 1 | 2 | 2 | 16 |
| which | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| where | 0 | 0 | 2 | 2 | 0 | 3 | 7 |
| when | 0 | 0 | 2 | 6 | 15 | 7 | 22 |
| how | 0 | 0 | 0 | 1 | 0 | 6 | 12 |
| why | 0 | 0 | 0 | 0 | 0 | 3 | 1 |
| Totals | 4 | 5 | 5 | 12 | 21 | 29 | 75 |

Looking at the relative clause totals, we can see that the construction "comes in" between 2.9 and 3.1, with differential control of the different subtypes.

But what does "come in" mean? It is a commonplace in the acquisition literature that different constructions "come in" at different points. But, again, what does "come in" mean?

It must mean that the licensing relation underlying the structure enters the grammar. Thus, for conjunction, it is the structural operation which takes two structures, conjoins them, and adds "and" shown in (1) and below. Note that, unlike Merge, this is a specific licensing relation, not a simple general building up relation of no specific type.

(5) Conjunction–licensing relation

$$\left.\begin{matrix} \text{s1} \\ \\ \text{s2} \end{matrix}\right| \longrightarrow \text{s1 and s2}$$

For relative clauses, it takes two structures and adjoins the second into the first.

(6) Relative clause adjunction–licensing relation

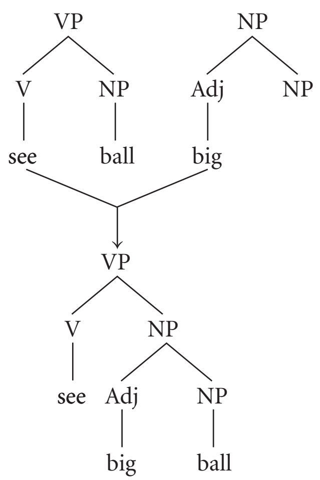
s1 NP V NP<br>
s2 CP $\longrightarrow$ (NP CP) V NP

Note that to say that a construction comes in, it must be possible to refer to the specific operation coming in. A general description, such as Merge, will not do, because it does not differentiate between the different types.

Each element in a tree is sanctioned in the tree by having some licensing relation sanction it in the tree. The tree is composed of these licensing relations (Appendix A).

For conjunction, it is not the word "and" which comes in. It is the licensing relation itself. For adjectival modification, it is not the words of the adjectives which "come in", it is the possibility of adjectival modification itself–which is a licensing relation. This is shown in (7). Note that the licensing relation is modeled by a generalized transformation.

(7) Adjectival modification–licensing relation (modeled by a generalized transformation)

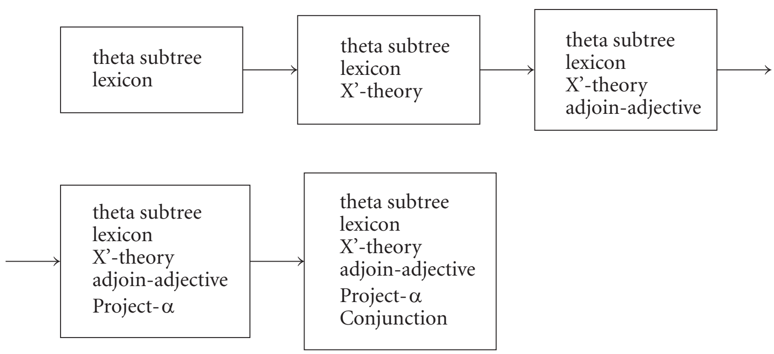



The above is the operation that may be said to enter the grammar when adjectives "come in". Note that it is not a general purpose building up operation like Merge, but instead it is characterized specifically by having the adjectival structure build up the tree. This, in turn, corresponds to the way very specific constructions mentioned above–relative clauses, conjunction–come into the grammar.

Different types of licensing relations are the following: conjunction, adjective adjunction, relative clause adjunction, and X'- theory. Each licensing relation builds up the tree by a generalized transformation, but note that the generalized transformation is not just Merge because it: i) may intersperse trees, ii) deals with syntactically meaningful units, and iii) includes the generalized transformation Project-α, which I will discuss later.

What does this mean for the child's grammar? This means that the additions to the child's grammar are the licensing relations, and the growth of the child's grammar is the addition of new licensing relations to it. I will assume that this takes place in a monotonic fashion. We may give a sketch of how the child's grammar grows.

(8)




The acquisition sequence may be summarized as follows:

(9) The acquisition sequence is the addition of licensing relations modelled by generalized transformations, added to the grammar.

Two comments. First, the licensing relations are representing different types of composition operations, not just Merge (which is basically exclusively bottom-up). They include operations which intersperse structure, as in Tree Adjoining Grammar, and they include as well the special operation Project-α, which I will discuss later. Second, if the lines of demarcation in the child's grammar are those shown above (adjoin-adjective, adjoin relative clause, conjoin, X'-theory, Project-α, etc.), then it constitutes an argument that Merge does not describe the adult grammar (at least not solely). The reason is that Merge is basically bottom-up, while the types here are of varying types. This view is that the grammar is based on relations and not on unstructured elements.

That is, we make the argument against Merge as a monolithic structure building operation on the following grounds.

(10) The lines of demarcation, shown in the acquisition sequence in (8), are the same as the lines of demarcation in the adult grammar. This means that different sorts of generalized transformations (adjunction, conjunction, X'-theory, etc.) put together the adult tree, not simply Merge as a single bottom-up operation.

To give a more exact idea of how licensing relations work, I will give them for one tree.

Let us suppose that the verb is gotten for free.

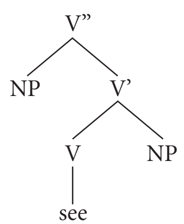
(11)


By X'-theory (the first licensing relation), this may be built up to V".

(12) X'-theory

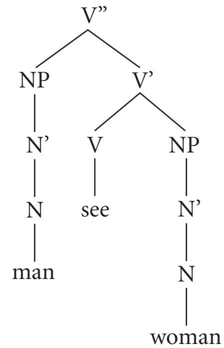

By enforcement of argument structure (the second licensing relation), the subject and the object may enter into the tree.

(13) Argument Structure



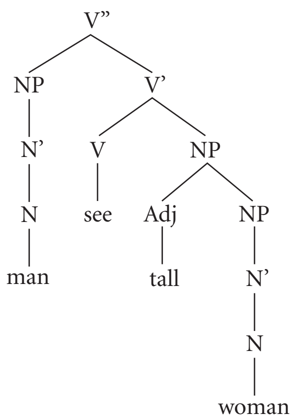

By X'-theory again, the tree may expand to the N, and lexical insertion is then given for free.

(14)



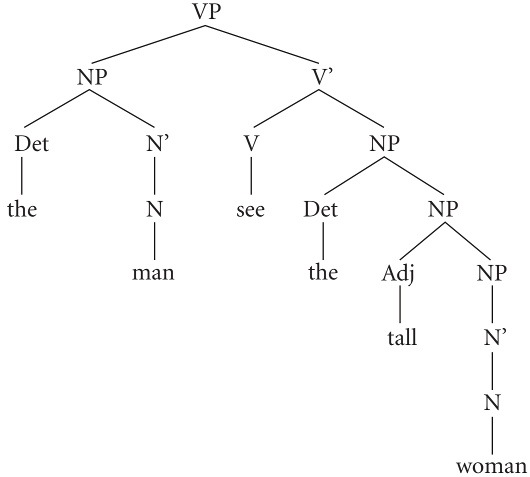

Adjoin-adjective, the third licensing relation, may now apply, adding in the adjective.

(15) Adjoin-adjective



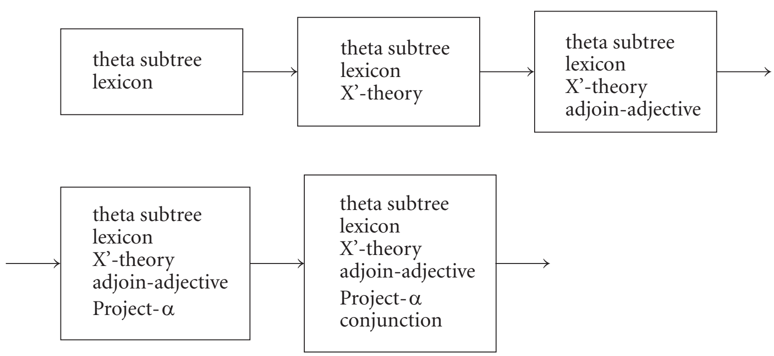

Finally, an operation I call Project-α, the fourth licensing condition, which I will discuss later, adds determiners into the representation.

(16) Project-α, fourth licensing relation




The claim then is that the child's grammar is composed of subsets of these licensing relations, gradually increasing over time. This is shown in (8) above, repeated here.

(17)




What is interesting is that if the child's grammar is described along these lines, with different licensing conditions giving descriptions at different stages, then it would be very strange that they be thrown out, and that simple Merge be used to describe the grammar as an adult. Rather, the adult grammar may be thought of as the composition of several of these more primitive relations.

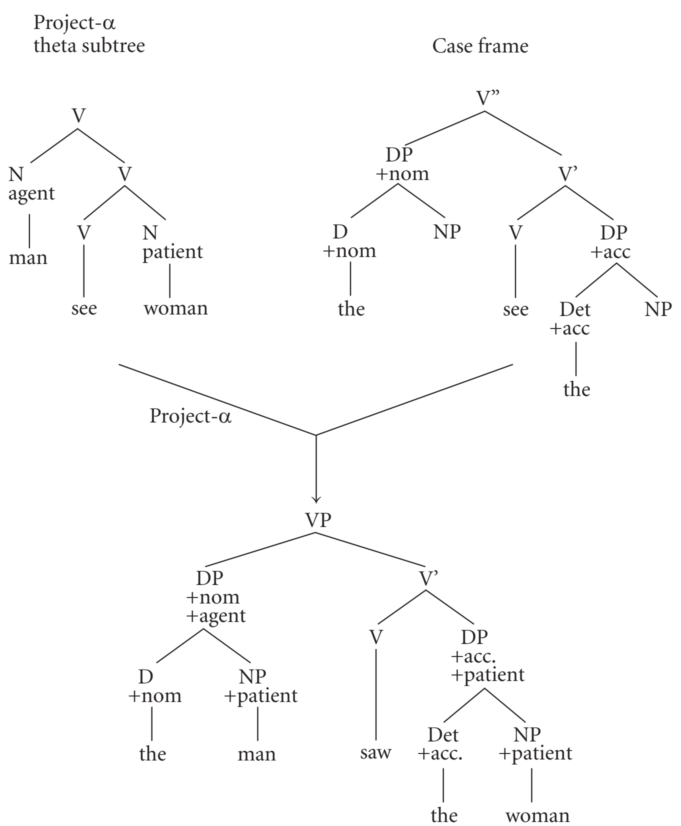
Let us add one final generalized transformation to our list of transformations. This is Project-α. Project-α takes trees composed of open class categories only, and projects them into a closed class frame (Lebeaux 1988, 2000a, b). The open class subtree is also described as being a pure representation of theta roles, while the closed class frame is a pure representation of Case. The Project-α transformation describes a key stage of child's grammar: namely, telegraphic speech. The open class tree describes the grammar of telegraphic speech, while the tree after the operation of Project-α describes the tree in the stage after telegraphic speech. Thus the pre- and post- Project-α stages correspond to the following stages of the child's grammar.

(18)

| set of licensing relations not including Project-α | → | set of licensing relations including Project-α |
|---|---|---|
| telegraphic speech | | post telegraphic speech |

I will assume that Project-α is a licensing relation which is like the others, yet it is special in that it takes the open class theta subtree, and closed class Case frame, and fuses the former into the latter. The open class elements in the theta subtree project into the open spaces in the Case frame.

(19)



Telegraphic speech exists at the level of the upper left quadrant above (the theta subtree). The Case frame, and the Post-Project-α grammar that gives rise to speech involving both open class and closed class elements.

Like the other generalized transformations described above (X' theory, adjective-adjunction, etc.), Project-α is intended to describe not simply the child's grammar, but the adult grammar as well. That is, in the derivational sequence of each sentence that an adult utters, the tree first exists as a theta subtree (the upper left quadrant above), and the Case frame is then constructed , and the theta subtree is projected into it. This is part of the adult derivation as well (Lebeaux 1988, 2000a, b). The other licensing relations I have given above similarly construct each adult tree.

In other work, I have given five types of evidence for the Project-α transformation. Here I would like to give just one of them. This can be found in the remarkable work on speech errors by Garrett (1975, 1980). Garrett, in this work and accompanying work, noted that a major class of speech errors had the following form: the closed class elements stayed rigid, while the open class elements permuted. Examples are given below. The permuted elements are underlined.

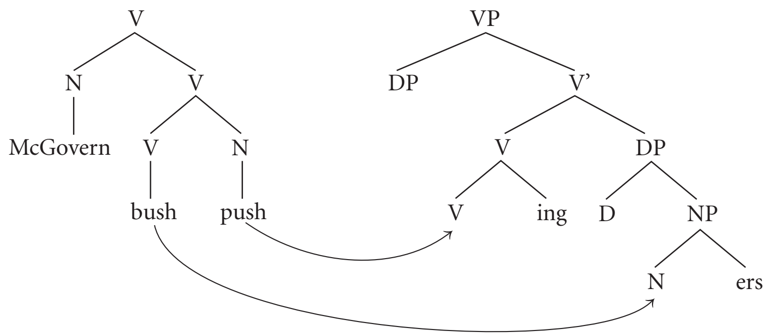
(20) Speech Error — Target

| Speech Error | | Target |
|---|---|---|
| my frozers are shoulden | —> | my shoulders are frozen |
| that's just a back trucking out | —> | that's just a truck backing out |
| McGovern favors pushing busters | —> | McGovern favors busting pushers |
| his sink is shipping | —> | his ship is sinking |
| the cancel has been practiced | —> | the practice has been cancelled |
| she's got her sets sight | —> | she's got her sights set |
| a puncture tiring device | —> | a tire puncturing device |

What is truly remarkable about these errors is that the closed class elements are fixed in place and the open class elements are permuted. A mistake occurs in the two levels of representation posited by Project-α and the result exchanges the open class element in two differing slots in the Case Frame. The permutation of elements is shown in (21) below, where the verb goes in the noun slot, and vice versa. That is (n push) goes into the verb slot and (v bust) goes into the noun slot.

(21) Project-α (speech error)
speech error: McGovern favors pushing busters

theta subtree | Case Frame




Permuted insertion:

a. (v bust) placed in nominal slot
b. (n push) placed in verbal slot
c. Closed class morphemes stay rigid

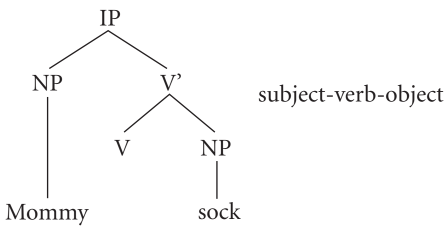
This pattern of switched open class insertion strongly suggests a grammatical module where all open class elements are in one form or tier and all closed class elements are in another tier, and the former are projected into the latter. It is only through such a grammatical formalism that these errors could be explained. Project-α is a rendition of such a formalism.

I would like to close by mentioning what I will call a Pattern of Naturalness in the child's grammar. Let us begin with a quote from Lois Bloom (1973)

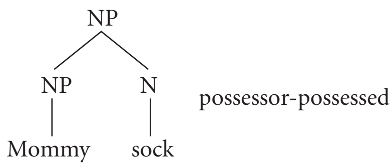
> However, more important reasons for arguing against the distributional evidence that would class "Mommy" as a pivot or functor form had to do with the fact that different utterances with "Mommy" meant different things. For example, in the first sample from Kathryn, the utterance "Mommy sock" occurred twice in two separate contexts.
>
> 1. Kathryn picking up Mother's sock.
> 2. Mother putting Kathryn's sock on Kathryn.
>
> It appeared that the difference in semantic interpretation between the two utterances (1) and (2) corresponded to a structural difference in grammatical relationship between the constituents "Mommy" and "sock". In one instance, the structure was a genitive relation and in the other the relation between subject and object.

The remainder of this section takes Bloom as a starting point to discuss naturalness and simplicity features in child's language. What does Bloom's comment mean, that these correspond to two different semantic relations? Put in other terms, Bloom is saying that there are primitive canonical binary relations, which hold over differing trees. One is a subject-verb-object tree, the other is a possessor-possessed tree

(22)




(23)



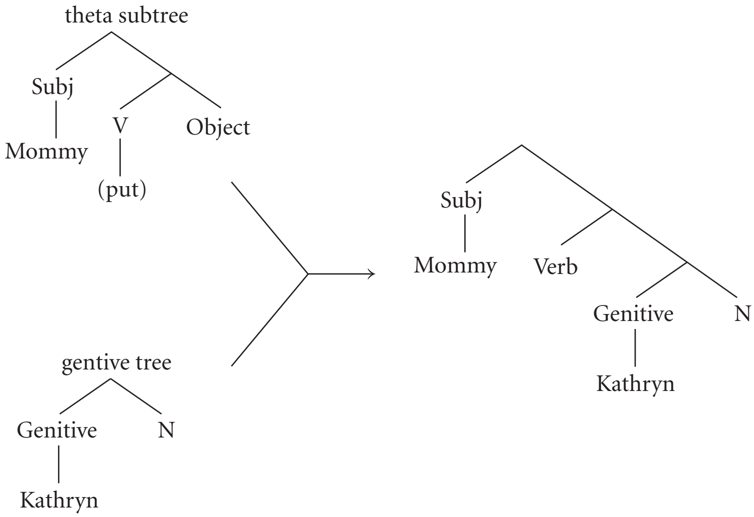

These structures are differentiated by, and crucially only by, the licensing relation which holds over the primitive tree. The Bloom example, is therefore much in line with, and supports, the argumentation here.

But let us now make a further point, which is actually the most interesting point about the early structures. Suppose that, corresponding to the structure in (22), the child had said instead of "Mommy sock", "Mommy Kathryn", meaning that Mommy was putting on Kathryn's sock.

(24) Situation: Mommy putting on Kathryn's sock
Output: Mommy Kathryn

Recall that we earlier had a theta subtree where the subject and object were part of a single initial tree. To get "Kathryn", a genitive modifier into the string, we need an operation which adjoins a genitive. This is what we need to drive "Kathryn" above into the tree. The generalized transformation takes the pure representation of argument structure and adjoins to it the genitive tree. This is shown in (25).

(25)



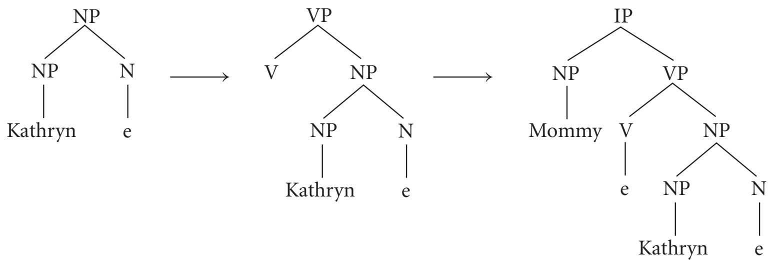

The child, instead of filling lexical entries in a single licensing tree, has to have two licensing trees, the theta subtree and the genitive, and compose them. Crucially "Mommy Kathryn" is therefore much more complicated in structure then "Mommy sock" in the licensing view here, because it involves putting together two separate trees.

The crucial theoretical point here is that "Mommy sock" would appear on a single theta subtree: hence it is natural (a natural utterance). "Mommy Kathryn" would not appear on a single licensing tree, but would require two trees which were adjoined, hence it is unnatural. That is, the licensing approach here marks "Mommy sock" as natural and "Mommy Kathryn" as unnatural. This is in accordance with their naturalness (and unnaturalness) in the actual output of the child. While the child often utters combinations like "Mommy sock", s/he rarely utters combinations like "Mommy Kathryn", where "Kathryn" is the genitive modifier of a nonexistent object NP. Therefore the licensing approach correctly captures the naturalness of the utterances.

(26) Mommy sock (natural)
Mommy Kathryn (unnatural)

However, in contrast, Merge cannot capture the naturalness of "Mommy sock" compared to the unnaturalness of "Mommy Kathryn". This is because it doesn't distinguish different grammatical primitives, but it is simply a binary build-up operation. The build up for "Mommy sock" (natural) and "Mommy Kathryn" (unnatural) would look almost identical for Merge. “Mommy Kathryn” would look as follows.

(27)




The Merge derivation in (27) is one in which the modifier modifies a null head, then the resultant is merged with a null verb, then the resultant with a subject. There is nothing non-natural or non-simple about this derivation. But this is the wrong result. For this derivation, which is natural for Merge is in fact unnatural for the grammar: sentences like “Mommy Kathryn” are much less common than “Mommy sock”. Merge does not distinguish between the two. But these are distinguished in the licensing approach, because “Mommy sock” are on the same basic theta tree, while “Mommy Kathryn” are on separate trees (and therefore an additional operation must apply to produce it).

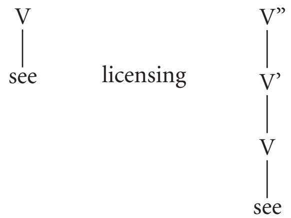
To conclude, I have argued that the language acquisition sequence consists of different licensing conditions entering the grammar one by one. These are constructions/operations like relative clause adjunction, conjunction, adjective adjunction, X’-theory, and Project-α. The acquisition sequence is “timed” by these constructions entering in. The licensing relations themselves are modeled by generalized transformations adding in new trees. Thus the hypothesis is that the grammar is based on relations, not on unstructured elements. The notion that a construction, for example, co-ordination, “comes in” is modeled by a generalized transformation which conjoins the elements; similarly, if the relative clause construction “comes in” at a certain time, this is modelled by an adjunction operation for relative clauses. The theory is therefore different than the theory of Merge, which has just a single building-up operation.

## Notes

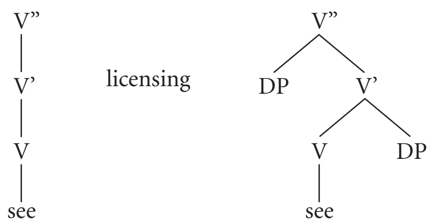
1. A much longer version of this paper may be found in Lebeaux 2000a.

## Appendix: Licensing relations

I am using a different method than Merge to build up the tree, namely licensing relations. The basic idea behind licensing relations is that given some element on the tree, some other element is sanctioned or licensed. Thus given a verb V, this licenses the build up to V". This is by X'-theory.

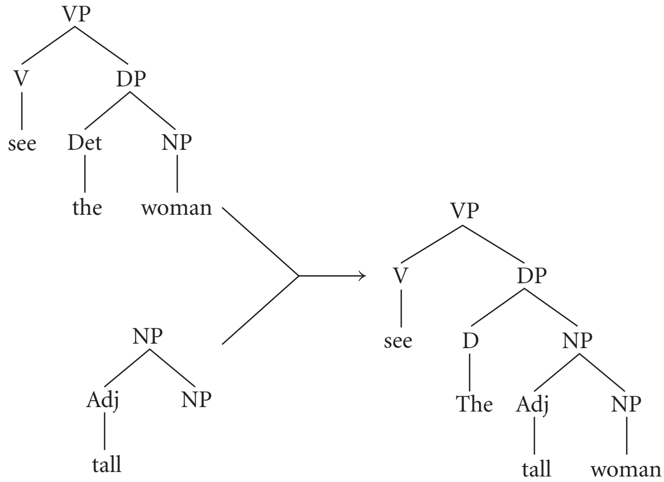
(1)



The argument structure of see licenses an object and a subject. Ultimately, the entire representation will be "roped" into the tree.




Four comments. First, while Abney, who perhaps introduced the term, held these to be part of the parser, I am taking these to be parts of the grammar itself. Second, the differing licensing relations are various in their type, operating in differing ways, unlike Merge which is simply building bottom-up. Third, unlike Abney, but like Lebeaux (2000a, b), I assume that the licensing relations do not hold in a stationary way over the trees, but are modeled often as generalized transformations. For example, the generalized transformation for adjective adjunction is the following:

(2)




Note, finally, that these operations may intersperse structure (unlike Merge).

The argument of this paper(also Lebeaux 2000a, b) is that such licensing relations give both the pattern of growth of the child and the organization of the derivation of the adult.

## References


Abney, S. (1987). Licensing and parsing. *Proceedings of NELS 17*.

Abney, S. & Cole, J. (1985). A Government Binding parser. *Proceedings of NELS, 15*.

Bloom, L. (1970). *Language Development: Form and Function in Emerging Grammars*, Cambridge, MA: MIT Press.

Bloom, L. (1973). The two-word stage reconsidered. In C. Ferguson and D. Slobin (Eds.), *Studies in Child Language Development* (pp. 430–440). New York: Holt, Rinehart, and Winston.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of Learning and Motivation, Vol. 9*, (pp. 133–177). New York: Academic Press.

Garrett, M. (1980). Levels of processing in speech processing. In B. Butterworth (Ed.), *Language Production, Vol. 1*, (pp. 177–220). New York: Academic Press.

Lebeaux, D. (1988). Language Acquisition and the Form of the Grammar. Ph.D. dissertation, University of Massachusetts.

Lebeaux, D. (2000a). A subgrammar approach to language acquisition. Technical Report 2000-077, NEC Research Institute.

Lebeaux, D. (2000b). *Language Acquisition and the Form of the Grammar*. Amsterdam: John Benjamins.

Miller, W. (1973). The acquisition of grammatical rules by children. In C. Ferguson and D. Slobin (Eds.), *Studies in Child Language Development* (pp. 380–391). New York: Holt, Rinehart, and Winston.

# Index

In the series LINGUISTIK AKTUELL/LINGUISTICS TODAY (LA) the following titles have been published thus far, or are scheduled for publication:

1. KLAPPENBACH, Ruth (1911-1977): *Studien zur Modernen Deutschen Lexikographie. Auswahl aus den Lexikographischen Arbeiten von Ruth Klappenbach, erweitert um drei Beiträge von Helene Malige-Klappenbach.* 1980.
2. EHLICH, Konrad & Jochen REHBEIN: *Augenkommunikation. Methodenreflexion und Beispielanalyse.* 1982.
3. ABRAHAM, Werner (ed.): *On the Formal Syntax of the Westgermania. Papers from the 3rd Groningen Grammar Talks (3e Groninger Grammatikgespräche), Groningen, January 1981.* 1983.
4. ABRAHAM, Werner & Sjaak De MEIJ (eds): *Topic, Focus and Configurationality. Papers from the 6th Groningen Grammar Talks, Groningen, 1984.* 1986.
5. GREWENDORF, Günther and Wolfgang STERNEFELD (eds): *Scrambling and Barriers.* 1990.
6. BHATT, Christa, Elisabeth LÖBEL and Claudia SCHMIDT (eds): *Syntactic Phrase Structure Phenomena in Noun Phrases and Sentences.* 1989.
7. ÅFARLI, Tor A.: *The Syntax of Norwegian Passive Constructions.* 1992.
8. FANSELOW, Gisbert (ed.): *The Parametrization of Universal Grammar.* 1993.
9. GELDEREN, Elly van: *The Rise of Functional Categories.* 1993.
10. CINQUE, Guglielmo and Guiliana GIUSTI (eds): *Advances in Roumanian Linguistics.* 1995.
11. LUTZ, Uli and Jürgen PAFEL (eds): *On Extraction and Extraposition in German*. 1995.
12. ABRAHAM, W., S. EPSTEIN, H. THRÁINSSON and C.J.W. ZWART (eds): *Minimal Ideas. Linguistic studies in the minimalist framework*. 1996.
13. ALEXIADOU Artemis and T. Alan HALL (eds): *Studies on Universal Grammar and Typological Variation.* 1997.
14. ANAGNOSTOPOULOU, Elena, Henk VAN RIEMSDIJK and Frans ZWARTS (eds): *Materials on Left Dislocation.* 1997.
15. ROHRBACHER, Bernhard Wolfgang: *Morphology-Driven Syntax. A theory of V to I raising and pro-drop.* 1999.
16. LIU, FENG-HSI: *Scope and Specificity*. 1997.
17. BEERMAN, Dorothee, David LEBLANC and Henk van RIEMSDIJK (eds): *Rightward Movement.* 1997.
18. ALEXIADOU, Artemis: *Adverb Placement. A case study in antisymmetric syntax.* 1997.
19. JOSEFSSON, Gunlög: *Minimal Words in a Minimal Syntax. Word formation in Swedish.* 1998.
20. LAENZLINGER, Christopher: *Comparative Studies in Word Order Variation. Adverbs, pronouns, and clause structure in Romance and Germanic*. 1998.
21. KLEIN, Henny: *Adverbs of Degree in Dutch and Related Languages.* 1998.
22. ALEXIADOU, Artemis and Chris WILDER (eds): *Possessors, Predicates and Movement in the Determiner Phrase.* 1998.
23. GIANNAKIDOU, Anastasia: *Polarity Sensitivity as (Non)Veridical Dependency.* 1998.
24. REBUSCHI, Georges and Laurice TULLER (eds): *The Grammar of Focus.* 1999.
25. FELSER, Claudia: *Verbal Complement Clauses. A minimalist study of direct perception constructions.* 1999.

26. ACKEMA, Peter: *Issues in Morphosyntax.* 1999.
27. RŮŽIČKA, Rudolf: *Control in Grammar and Pragmatics. A cross-linguistic study.* 1999.
28. HERMANS, Ben and Marc van OOSTENDORP (eds): *The Derivational Residue in Phonological Optimality Theory*. 1999.
29. MIYAMOTO, Tadao: *The Light Verb Construction in Japanese. The role of the verbal noun.* 1999.
30. BEUKEMA, Frits and Marcel den DIKKEN (eds): *Clitic Phenomena in European Languages.* 2000.
31. SVENONIUS, Peter (ed.): *The Derivation of VO and OV*. 2000.
32. ALEXIADOU, Artemis, Paul LAW, André MEINUNGER and Chris WILDER (eds): *The Syntax of Relative Clauses.* 2000.
33. PUSKÁS, Genoveva: *Word Order in Hungarian. The syntax of $\bar{A}$-positions.* 2000.
34. REULAND, Eric (ed.): *Arguments and Case. Explaining Burzio's Generalization.* 2000.
35. HRÓARSDÓTTIR, Thorbjörg. *Word Order Change in Icelandic. From OV to VO.* 2000.
36. GERLACH, Birgit and Janet GRIJZENHOUT (eds): *Clitics in Phonology, Morphology and Syntax.* 2000.
37. LUTZ, Uli, Gereon MÜLLER and Arnim von STECHOW (eds): *Wh-Scope Marking.* 2000.
38. MEINUNGER, André: *Syntactic Aspects of Topic and Comment.* 2000.
39. GELDEREN, Elly van: *A History of English Reflexive Pronouns. Person, ''Self'', and Interpretability.* 2000.
40. HOEKSEMA, Jack, Hotze RULLMANN, Victor SANCHEZ-VALENCIA and Ton van der WOUDEN (eds): *Perspectives on Negation and Polarity Items.* 2001.
41. ZELLER, Jochen : *Particle Verbs and Local Domains.* 2001.
42. ALEXIADOU, Artemis : *Functional Structure in Nominals. Nominalization and ergativity.* 2001.
43. FEATHERSTON, Sam: *Empty Categories in Sentence Processing.* 2001.
44. TAYLAN, Eser E. (ed.): *The Verb in Turkish.* 2002.
45. ABRAHAM, Werner and C. Jan-Wouter ZWART (eds): *Issues in Formal German(ic) Typology.* 2002.
46. PANAGIOTIDIS, Phoevos: *Pronouns, Clitics and Empty Nouns. 'Pronominality' and licensing in syntax.* 2002.
47. BARBIERS, Sjef, Frits BEUKEMA and Wim van der WURFF (eds): *Modality and its Interaction with the Verbal System.* 2002.
48. ALEXIADOU, Artemis, Elena ANAGNOSTOPOULOU, Sjef BARBIERS and Hans-Martin GAERTNER (eds): *Dimensions of Movement. From features to remnants.* 2002
49. ALEXIADOU, Artemis (ed.): *Theoretical Approaches to Universals.* 2002.
50. STEINBACH, Markus: *Middle Voice. A comparative study in the syntax-semantics interface of German.* 2002.
51. GERLACH, Birgit: *Clitics between Syntax and Lexicon.* 2002.
52. SIMON, Horst J. and Heike WIESE (eds): *Pronouns – Grammar and Representation.* 2002.

53. ZWART, C. Jan-Wouter and Werner ABRAHAM (eds): *Studies in Comparative Germanic Syntax. Proceedings from the 15th Workshop on Comparative Germanic Syntax (Groningen, May 26-27, 2000)(Workshop).* 2002.
54. BAPTISTA, Marlyse: *The Syntax of Cape Verdean Creole. The Sotavento varieties.* 2002.
55. COENE, M. and Yves D'HULST (eds): *From NP to DP. Volume 1: The syntax and semantics of noun phrases.* n.y.p.
56. COENE, M. and Yves D'HULST (eds.): *From NP to DP. Volume 2: The expression of possession in noun phrases.* n.y.p.
57. DI SCIULLO, Anna-Maria (ed.): *Asymmetry in Grammar. Volume 1: Syntax and semantics.* 2003.
58. DI SCIULLO, Anna-Maria (ed.): *Asymmetry in Grammar. Volume 2: Morphology, phonology, acquisition.* 2003.
59. DEHÉ, Nicole: *Particle Verbs in English. Syntax, information structure and intonation.* 2002.
60. TRIPS, Carola: *From OV to VO in Early Middle English.* 2002.
61. SCHWABE, Kerstin and Susanne WINKLER (eds.): *The Interfaces. Deriving and interpreting omitted structures.* 2003.
62. CARNIE, Andrew, Heidi HARLEY and Mary WILLIE (eds.): *Formal Approaches to Function in Grammar. In honor of Eloise Jelinek.* 2003.