Phonology **15** (1998) 317–352. Printed in the United Kingdom

*Head–dependent asymmetries in phonology: complexity and visibility**

B. Elan Dresher
University of Toronto

Harry van der Hulst
Holland Institute of Generative Linguistics and University of Leiden

0 Introduction

Developments in phonological theory have led to the recognition that phonological representations have a layered constituent structure. Many, perhaps all, of these constituents contain elements which can be identified as HEADS. Heads enter into various kinds of relations with their dependents. In this article, we identify a phenomenon which is quite pervasive in every part of phonology which has heads and dependents, namely, the existence of HEAD–DEPENDENT ASYMMETRIES (henceforth HDAs). While various particular manifestations of these asymmetries are well known and have been much studied, this is the first attempt, to our knowledge, to unite a broad range of seemingly different phenomena under one heading. We identify various types of HDAs, and propose constraints on possible HDAs.

Most importantly, we distinguish between HDAs that involve COM-

* Work on this article began when the coauthors collaborated in the Project on Language Acquisition sponsored by NIAS (Netherlands Institute for Advanced Study in the Humanities and Social Sciences) in 1990–91. Early versions were presented at: the GLOW Workshop on Acquisition of Phonology, Leiden University, March 1991; the Workshop on Formal Issues in Language Acquisition, Max-Planck-Institut für Psycholinguistik, Nijmegen, June 1991; the MOT Phonology Workshop, University of Ottawa, January 1992; and the 1st HIL Conference on Phonology, Leiden University, January 1993. We are grateful to participants at these workshops, to colleagues and students at Leiden and Toronto, and to several anonymous reviewers for valuable comments. This article is a considerably revised version of Dresher & van der Hulst (1993, 1995). In addition to the organisations mentioned above, Dresher acknowledges the support of SSHRCC (Social Sciences and Humanities Research Council of Canada) Research Grants 410-92-0885 and 410-96-0842.

The above-mentioned NIAS project was organised by van der Hulst together with Teun Hoekstra, to whose memory we dedicate this article.

plexity, and those that involve visibility. These have properties which potentially contradict each other. We propose that they apply in fundamentally different types of cases: unlike complexity HDAs, visibility HDAs are limited to mappings from one phonological plane to another, and so are related to the notion of projection (cf. Vergnaud 1977).

We also wish to show that an understanding of HDAs reveals general structural principles that play a role in diverse phenomena at various levels of the phonological hierarchy. For example, the fact noted in the Optimality Theory literature that certain positions tend to be more 'faithful' to underlying specifications (Beckman 1998) is a consequence of the fact that heads allow more complexity. These principles act as constraints on possible constraints, and on possible mappings from one plane to another.

Our proposal will include the following claims.

First, there is a meaningful notion of phonological head that cuts across segmental and suprasegmental levels, indeed perhaps across phonology and syntax; cf. Anderson & Ewen (1987: 283f) and Halle & Vergnaud (1987: 6), who assume that constituency and heads are fundamental linguistic concepts.

Second, phonological heads show the maximum complexity allowed by a grammar. Thus, heads and dependents may be equally complex; but if there is an asymmetry, it will always be the head that is more complex than the dependent. Any asymmetry of this type is an HDA. We will discuss a number of these HDAs at various levels of the phonology above the level of the syllable. We then also consider HDAs in the phonology of sign languages, showing that certain asymmetries in two-handed signs form clear illustrations of heads allowing greater complexity than dependents.

Third, and of particular interest to us in this article, we identify a distinct class of HDA in which heads, contrary to our hypothesis, appear to be less complex than their dependents. We will argue that these involve two types of cases. In the first type, there is no true asymmetry of complexity, but rather of visibility: in such cases, constituent units of a head are visible to a greater depth than the corresponding constituents of a dependent with respect to some mapping or projection of phonological structure from one plane to another. The main focus of this article is how we can distinguish between complexity and visibility HDAs. The second type of case involves real violations of the HDA principle. However, we argue that such violations do not occur arbitrarily, but result from another special property of heads, and involve different types of projections for heads and dependents.

Fourth, we locate the origin of HDAs in the process of acquisition.

This article is organised as follows. In §1 we consider in some detail terms such as complexity, visibility and dependency, and identify various kinds of complexity HDAs. In §2, we present some illustrative examples of HDAs that involve visibility at the higher levels of the prosodic hierarchy. In §3, we discuss the differences between complexity and visibility HDAs, and propose criteria for characterising the situations in which each might be expected to exist. In §4, we consider some further visibility asymmetries at the word level, including some that go the wrong way, and identify the source of these 'anti-HDAs'. In §5 we briefly consider two further topics relevant to HDAs: in §5.1 we consider how HDAs play a role in the phonology of signs, and §5.2 has a brief discussion of the connection between HDAs and language acquisition. §6 is a conclusion.

1 Complexity asymmetries

The claim that heads are or may be more complex than dependents is quite commonly encountered in linguistic discussions, or at least it is one of those claims that most linguists immediately agree with.[1] In this section we wish to make it clear that the form in which the claim is usually stated obscures the fact that the relevant asymmetry can be manifested in a variety of ways.[2] Our goal in this section is to make explicit the various ways in which the complexity of heads exceeds that of dependents.

Let us note at the outset that complexity is a relative notion, so when we say that a constituent is complex, we always mean relative to one that is less complex. With that understanding, we will first make a distinction between two types of complexity (local and non-local) in §1.1. §1.2 provides illustrations of this distinction with reference to the foot level. §1.3 explains the distinction between paradigmatic and syntagmatic asymmetries. At the end of this section (in §1.4) we make explicit that the whole discussion so far has been limited to what we will call α-α head–dependent relations (where head and dependent are the same type of constituent). We distinguish these from α-β relations, which fall outside the scope of our discussion. §1.5 addresses the relation between complexity and phonetic exponency. In §1.6 we summarise our findings.

1.1 Local and non-local complexity

We will say that a node C has LOCAL COMPLEXITY if it branches, whereas other nodes of the same type do not branch, as in (1a); or a node can be

[1] One of the few to make this claim explicit is Harris (1990: 274), in the form of the Complexity Condition, which requires that a governee may be no more complex than its governor. This is a particular form (syntagmatic) of head–dependent asymmetry.

[2] A version of this idea has been the subject of a growing literature in Optimality Theory in terms of increased faithfulness in certain positions. McCarthy & Prince (1995: 364) propose a metaconstraint to the effect that faithfulness in roots is universally ranked higher than faithfulness in affixes. Thus, if a marked feature is allowed to surface in an affix, it must also be allowed to surface in a root, but not *vice versa*. Similarly, Beckman (1997: 13) proposes that the root-initial syllable in Shona is singled out, a 'higher premium being placed on the maintenance of underlying contrasts in initial syllables than elsewhere in the word'. In other words, certain features can surface in the initial syllable that may be suppressed in other syllables. Why are certain positions singled out for increased faithfulness? In our terms, it is because they are heads. As we will show, head–dependent asymmetries involve more than increased faithfulness.

locally complex relative to other nodes if it has an immediate dependent when others do not (1b):

(1) *Local HDAs*

a. *Complex* *Simple* b. *Complex* *Simple*

Some instantiations of these types of asymmetries are given in (2). Example (2a) illustrates the HDA in (1a), where a prosodic word that comprises two metrical feet is more complex than a word which has only one foot. Example (2b) is of the type (1b): in this instance, a CORONAL place node that has a dependent is more complex than one which does not:

(2) *Instantiations of local HDAs*

a. *Complex* *Simple* b. *Complex* *Simple*

A second type of complexity relevant to our discussion is NON-LOCAL COMPLEXITY, which refers to the internal structure of the immediate daughter of a node C; as we will see, this daughter is the head of C. In (3a) and (3b), for example, node C is non-branching, and so has the same amount of local complexity, in both cases; however, there is a difference in the complexity of node D:

(3) *Non-local complexity*

a. *Complex* b. *Simple*

An example of a non-local HDA is given in (4): a prosodic word comprising a single foot that itself dominates two moras is more non-locally complex than a word that dominates a monomoraic foot:

(4) *Instantiation of non-local HDA*

a. *Complex* b. *Simple*

We will call asymmetries such as in (1) and (3) COMPLEXITY HDAs (HDA-Cs). Our claim is that HDA-Cs can only go in one direction: heads may tolerate more complexity than dependents, but the reverse situation is not permitted.[3]

In the next section, we will illustrate the distinction between local and non-local asymmetries with reference to foot structure.

1.2 Local and non-local foot-levels HDAs

Foot assignment may be sensitive to aspects of syllable structure that are local, in that geometric properties of the syllable rhyme are relevant, or non-local, in the sense that segmental information appears to be relevant. Hayes (1995) separates QUANTITY, which involves syllabic complexity (i.e. long *vs.* short vowels, closed *vs.* open syllables), from PROMINENCE, which involves segmental properties; in our terms, this is just the distinction between local and non-local complexity. We believe that viewing the distinction in this way opens new perspectives on the analyses of various stress systems in which the factors determining weight appear to be somewhat opaque, or a 'mixture' of various factors.[4]

With respect to local complexity asymmetries it has been noted that a model in which feet consist of syllables faces a conceptual problem if the latter consist of an onset and rhyme constituent. The problem is that the asymmetries are not really local, because the syllable node intervenes between the foot terminal and the (relevant) rhyme node. This problem has led to two responses. One has been to deny the existence of the syllable node. Feet are built on rhymes, while onsets are skipped or left unparsed. This is the position that we find in metrical analyses that adopt the government phonology approach to 'syllabic' structure (cf. Kaye *et al.* 1990). The other response involves abandoning the onset–rhyme theory in favour of a moraic theory in the style of Hyman (1985) or Zec (1988), in which 'onset' material is dominated by the first mora.[5] In §4.1 we will

[3] It should be apparent from the examples in (2) and (4) that notions of complexity are themselves quite theory-bound. For example, in theories with binary and non-privative place features, the different segments in (2b) will not be representable as differing in structural complexity. Such theories may have other ways of encoding what we call complexity differences: for example, one could have a markedness scale that stipulates relative markedness. In such a theory, HDAs would still exist, but they would be evaluated with respect to the scale of markedness. Cases such as (2a), however, would presumably be represented as complexity differences in all theories. In this article, we will assume that all such differences can be represented in terms of structural complexity.

[4] It is therefore useful to reserve the term SYLLABLE WEIGHT to distinguish between heavy and light syllables, however they are characterised in a specific language. Thus, rhymal or nuclear complexity (including quantity) is one expression of weight (local complexity); non-local weight distinctions could include segmental properties.

[5] This issue is discussed in van der Hulst (1994), where it is observed that the mora division posited by Hyman (1985) fits better in the prosodic hierarchy, where every level n is divided into uniform units of level $n-1$; moras fit the uniformity criterion better than the disparate units of onset and rhyme.

argue that feet are always built directly on moras. We do not see moras as constituents of the syllable, however, but as units of weight that are projected from syllable constituents (following Halle & Vergnaud 1987).

We will now discuss a particularly common type of case in which the distinction between head and dependent inventories involves non-local complexity, i.e. complexity at the segmental level. Rather than focusing on prominence effects (as characterised above), we will look at a vowel-neutralisation phenomenon.

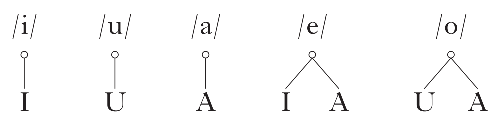
As an example, consider certain dialects of Modern Greek (Newton 1972), in which strong syllables may contain any of the five vowels in (5a), whereas weak syllables can contain only those in (5b):

(5) *Greek vowels*

a. *In strong syllables*	b. *In weak syllables*
i u	i u
e o	
a	a

In order to be able to classify this as a case of a non-local complexity asymmetry, we need a theory of segmental representation which ranks the three peripheral vowels as less complex than the two mid vowels. Most theories which view vowels as being composed of combinations of basic particles or elements – versions of dependency phonology (Anderson & Ewen 1987), particle phonology (Schane 1984) and government phonology (Kaye *et al*. 1985) – have this property, as shown in the minimal representations in (6):

(6) *Vowel structure*



In such theories, the basic units are unary features or 'elements' that can occur separately or in combination (for further development of this idea see van der Hulst 1994). As shown, mid vowels are branching under this view, and more complex than the other vowels. In the above example, then, the generalisation is that vowels with branching structures are permitted only in stressed syllables.

Asymmetries between sets of stressed vowels and sets of unstressed vowels having roughly this character are quite common. A number of such cases were collected by Trubetzkoy (1969). Trubetzkoy's interest was in developing a theory of neutralisation, which has obvious points of contact with our approach. Thus, in Trubetzkoy's terms, privative and gradual oppositions involve a contrast between an unmarked segment and a segment or segments which are marked for the property in question. Since the neutralisation product of such oppositions (the archiphoneme) 'can only contain that which is common to both opposition members' (1969: 82), it follows that they will neutralise to the unmarked member. Unary feature theories claim that segmental complexity will reflect markedness: if asymmetries between stressed and unstressed vowel sets reflect complexity restrictions on the latter, we expect that the more complex vowels will be the ones that are missing.[6]

In this section we have focused on foot-level asymmetries. We have given examples of both local HDAs, involving syllable quantity, and non-local HDAs, involving prominence effects and vowel neutralisation. In the following sections we will discuss examples that involve asymmetries holding at higher levels of the prosodic hierarchy. At levels of the hierarchy lower than the foot (i.e. involving the syllable and the segment), we face fundamental differences of opinion regarding the nature of the units that are combined to form constituents. At the syllabic level, as we have already seen, we are confronted with two rather different views: moraic theory and onset–rhyme theory. Similarly, different views of segment structure have different implications for evaluating complexity, and hence for HDAs. To pursue the study of HDAs at these levels, we therefore first have to decide what these units are before we can investigate the way in which HDAs manifest themselves. In this article we will not discuss HDAs at the syllable and segmental levels. For discussion of HDAs at the segmental level we refer to van de Weijer (1994) and van der Hulst (1994, 1996a).[7]

In the next sections we identify some further distinctions relevant to the formal characteristics of HDAs.

1.3 Paradigmatic and syntagmatic HDAs

To say that the head position (H) allows a greater complexity than a dependent position (D) conflates two different types of asymmetry:

(7) a. *Paradigmatic* (*lexical*) *asymmetry*
The set of units, S_H, that may occur in H and the set of units, S_D, that may occur in D is non-identical: $S_H \neq S_D$.

b. *Syntagmatic asymmetry*
In a specific combination, C, consisting of a head, H, and one or more dependents, D, what occurs in D may not be more complex than what occurs in H: $C_H \geqslant C_D$.

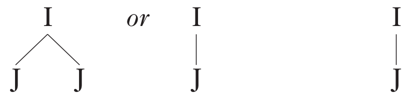
[6] See Dresher & van der Hulst (1993, 1995) for further discussion of these types of cases.

[7] Van de Weijer (1994) discusses HDAs on the segmental level in the context of a dependency-based model.

A paradigmatic local asymmetry is illustrated in (8); heads may branch or not, but dependents may not branch:

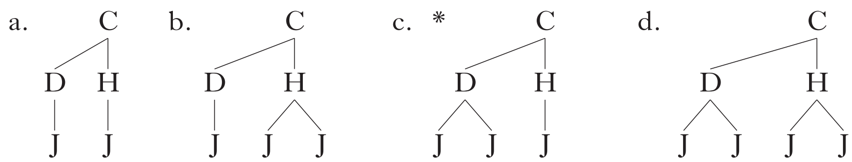
(8) *Paradigmatic HDA*

a. *Heads* b. *Dependents*



A syntagmatic local asymmetry is illustrated in (9):

(9) *Syntagmatic HDA: permissible and impermissible constituents*



In this example, heads and dependents may both branch, so there is no paradigmatic asymmetry; however, in any given structure, the dependent (D) may not be more complex than the head (H). Thus, (9c) is ruled out. The Greek example above illustrates a paradigmatic HDA: the complex vowels /e/ and /o/ can appear in a stressed syllable, but not in an unstressed syllable. A syntagmatic version of this HDA would occur in a language in which mid vowels may occur in an unstressed syllable, but not when the stressed vowel is high.[8]

The difference between paradigmatic and syntagmatic HDAs can also be illustrated with reference to foot-level asymmetries. In a language with a quantity-insensitive (QI) stress system, syllable weight is not evaluated for purposes of the metrical system. In such a system, the internal structure of the syllable is not a relevant property for evaluating complexity; any syllable can occupy any position in a foot without incurring a complexity asymmetry. However, in a quantity-sensitive (QS) system, weight is relevant, and we have to recognise a distinction between heavy (H) and light (L) syllables.

In foot theory, quantity-sensitivity comes in a number of types, given as (10a–d). For each type, we show the feet which are permitted (we assume syllabic left-headed binary (trochaic) feet):

[8] Dyck (1995) proposes that metaphony in some Spanish and Italian dialects has the effect of eliminating syntagmatic HDA violations at the segmental level.

(10) *Types of quantity-sensitive trochees*

Type	Ḣ L	Ĺ L	Ḣ H	Ĺ H	*Type of HDA*
a. No-mismatch	*yes*	*yes*	*yes*	*no*	syntagmatic
b. QS	*yes*	*yes*	*no*	*no*	weak paradigmatic
c. Revised Obligatory Branching	*yes*	*no*	*yes*	*no*	weak paradigmatic
d. Obligatory Branching	*yes*	*no*	*no*	*no*	strong paradigmatic

(10) lists all logical possibilities which do not violate the HDA principle. These foot types fall on different points of a path from minimal to maximal quantity-sensitivity. Minimal QS results in the non-mismatch foot (10a): the only situation that is avoided in this case is one in which the intrinsic weight of the syllables clashes with the extrinsic prominence relation imposed by foot organisation. No-mismatch feet are discussed by van der Hulst (1984), and figure in the analysis of Finnish secondary stress of Grijzenhout (1992).

The other pole, maximal QS, results in Obligatory Branching (OB) feet (10d), introduced in Vergnaud & Halle (1979) and Hayes (1980). Here, the very presence of a foot requires a heavy syllable, whereas the dependent position only supports a light syllable.

There are two less extreme possibilities. In foot type (10b), [Ḣ H] feet are disallowed. This is generally considered to be the standard form of QS: in head position, both heavy and light syllables may occur, but in dependent position only light syllables are tolerated. Alternatively, Hammond (1986) proposes the Revised Obligatory Branching (ROB) type in (10c), which allows [Ḣ H], but bans [Ĺ L].

The differences between these foot types exemplify the distinction between paradigmatic and syntagmatic HDAs. Recall that a syntagmatic relation holds between the sister nodes that form an actual constituent, whereas a paradigmatic relation holds between the sets of units that may occur in the head or dependent position.

A syntagmatic asymmetry, wherein a given foot has a head which is less complex than its dependent, i.e. [Ĺ H], is ruled out by all QS foot types. All QS foot types conform to the syntagmatic HDA. With reference to (10a) this is all that needs to be said. There is no paradigmatic asymmetry in (10a), since the set of syllables which can appear in the head (S_H) is the same as the set of syllables which can appear in the dependent (S_D): in both cases, the set comprises {H, L}. In the case of OB feet (10d), S_H and S_D are disjoint (non-overlapping): $S_H = \{H\}$ and $S_D = \{L\}$. We refer to this kind of paradigmatic asymmetry as 'strong', because there is no intersection between the head and dependent lexicon. The other two cases instantiate relations in which $S_D = \{L\}$ is properly included in $S_H = \{L, H\}$ or *vice versa*: $S_H = \{H\}$ is properly included in $S_D = \{L, H\}$ (ROB). We call this a weak paradigmatic asymmetry. Both cases anticipate a paradigmatic asymmetry by restricting either the 'lexicon of heads' (S_H) or the 'lexicon of dependents' (S_D).

The ROB foot is especially interesting because it shows that it is not necessarily the case that the head position always reveals the full range of possibilities that a language allows. For this reason, it is incorrect to say that head positions are positions of maximum *contrast*. The correct formulation, in our view, is that head positions are positions of maximum *complexity*. Where heads and dependents have the same *lower* limits of complexity (11a), then it is the case that head positions will allow for the maximum contrasts. But it is also possible that the lower limit on complexity in the head position is set higher than in the dependents (11b); in this case, head positions allow for more contrasts at the upper end of the complexity scale, whereas dependent positions allow for more contrasts at the lower end. Finally, it is also possible to have a situation as in (11c), where heads and dependents have the same upper bound on complexity, but where heads have a higher lower bound. In such cases, dependent positions will display a greater number of contrasts than heads.

(11) *Upper and lower bounds on complexity in heads and dependents*

a. H	D	b. H	D	c. H	D
—		—		—	—
	—		—		
		—		—	
—	—		—		—

1.4 α-α *vs.* α-β dependency

Head–dependent relations can be of two kinds: a head and a dependent can be constituents of the same type, or they can be constituents of different types. Examples of the first kind, which we call α-α DEPENDENCY, are a series of feet in a word, of which one is the head and the others are dependents; or a series of syllables in a foot, or phonological phrases in an intonational phrase, and so on. The second kind, which we call α-β DEPENDENCY, involves heterogeneous constituents: for example, the relation between a nucleus and a coda, or a rhyme and its onset. An example at the segmental level would be manner and place nodes, if these enter into a dependency relation (van der Hulst 1994, 1996a). Our discussion of HDAs is limited to α-α dependencies, for the simple reason that the potential complexity of heads and dependents in such relations is comparable. When the constituents are of different types, the complexity measures relevant to each may be different.[9]

[9] This is not to say that there may not be relevant complexity relations between head and dependent in α-β dependencies: for example, there may be implicational rules holding of the branching allowed of an onset *vis-à-vis* a rhyme, but we do not discuss these here. Harris (1990), developing ideas of Kaye *et al*. (1985, 1990), proposes that a segment which is in a governing position may not be less complex than a segment in the position it governs, where the positions may be within a syllable constituent (within an onset or nucleus, for example) or across constituents (an onset governing a preceding rhymal complement). In our terms, these are syntagmatic HDAs.

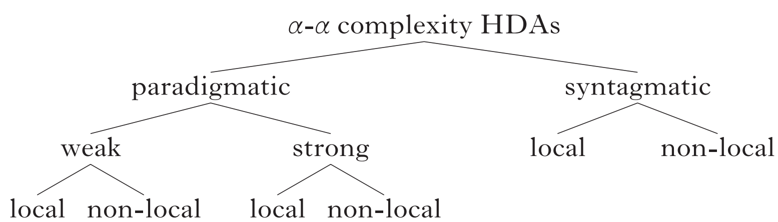
1.5 Relevant properties

We should stress that, when we refer to heads and dependents, we have in mind units that are defined with respect to prosodic or metrical structures that exist in a given language. That is, a head is not only a constituent that is somehow phonetically more prominent or salient than a dependent, but must have a well-defined relation to that dependent in terms of some higher-level structure.[10] It is also necessary to bear in mind that not every property of a phonological unit is relevant to computing its complexity for purposes of HDAs. For example, in a quantity-insensitive stress system, the internal structure of a syllable has no bearing on its complexity as either a head or dependent of a metrical foot: every syllable in such a stress system has equal weight, regardless of its internal composition. By contrast, in a quantity-sensitive stress system, some aspects of the internal structure of syllables are relevant in determining the weight of a syllable; these properties, which may vary to some extent from language to language, are thus relevant in assessing the complexity of the constituents of foot structures, while other properties are not.

1.6 Summary

In this section we have established that head–dependent complexity asymmetries can be of different types, as shown in (12):

(12)



The typology in (12) in principle recurs at every level of the prosodic hierarchy. Due to space limitations, it is beyond the scope of the present article to provide illustrations of all the logically possible cases. Moreover, the primary concerns of this article lie elsewhere, i.e. with how we should account for situations that seem to contradict the validity of the direction of the asymmetry, i.e. cases in which the head is seemingly less complex than its dependent(s). §§2 and 3 focus on another type of HDA. It will be claimed that the asymmetries found here do not involve relations between

[10] Here, then, we take a position that is fundamentally different from that presented in Steriade (1994) and Beckman (1997), who characterise positions only in terms of their phonetic salience. Without denying that heads may be phonetically salient, we observe that head positions play a special role in the phonology as well as the phonetics, and that phonetic salience may be a reflection of phonological structure.

sister nodes within a hierarchical organisation, but rather the projection of information from one plane to another.

2 Visibility asymmetries

We will now consider another type of asymmetry between heads and dependents which appears to be rather widespread. We will show that this type of asymmetry is not an asymmetry of *complexity*, but of *visibility* or *accessibility*. We will illustrate this type of asymmetry with two cases, which we will consider in some detail. Both concern the organisation of words into larger prosodic constituents. The first case concerns asymmetrical weight effects in Tiberian Hebrew, and the second case is an example of a somewhat similar asymmetry in Ligurian Italian.

We assume the prosodic hierarchy in (13):[11]

(13) *Prosodic hierarchy*

I	intonational phrase
|	
ϕ	phonological phrase
|	
ω	phonological word
|	
F	foot
|	
…	

2.1 Tiberian (Biblical) Hebrew

Evidence for Biblical Hebrew prosodic structure comes from the elaborate notation developed over several centuries until ca. 900 by scholars living near the city of Tiberias. That this notation serves as a reasonably reliable transcription of Biblical Hebrew prosody is argued by Dresher (1994); here we will assume the correctness of this interpretation, and focus on some properties of the prosodic representation.

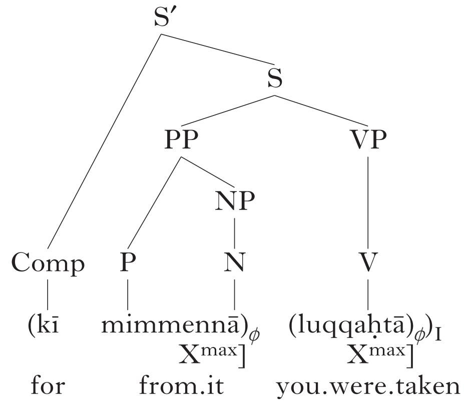
In general, Biblical Hebrew phonological phrases are delimited by the right edges of syntactic categories. We can observe the effects of mapping to the right edge of X^{max} in examples like (14a), where the PP, which consists of a single prosodic word, forms a phrase together with the preceding complementiser (cutting across a syntactic boundary), leaving the verb to form another phrase. In (14b), a phrase boundary again occurs

[11] We take no stand here on whether there are additional levels above the word, such as the clitic group (Hayes 1989) or the minor and major phrase (Selkirk 1986).

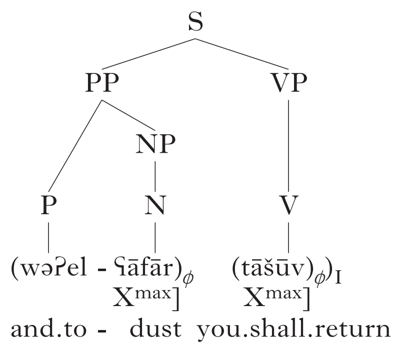
between the PP and the V (with cliticisation of the preposition to the following noun).

(14) *Syntax and phrasing:* X^{max} right

a. *Gen.* 3.19: *End of half verse*



b. *Gen.* 3.19: *End of verse*



When a verb or noun has a coordinate complement immediately to its right, the phrasing usually groups the first member of the compound with the preceding governor, as in (15). This type of phrasing is consistent with putting a phrase boundary at the right edge of X^{max}: the first such edge in (15) coincides with the end of the first NP of the complement.

(15) *Phrasing of X* [*Y and-Y*] (*X governs Y*)

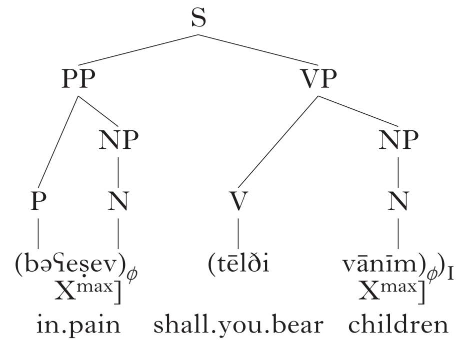
(kabbēð ʔeθ-ʔāvīxā) (wəʔeθ-ʔimmexā)

honour ACC.your.father and.ACC-your.mother (*Deut.* 5.16)

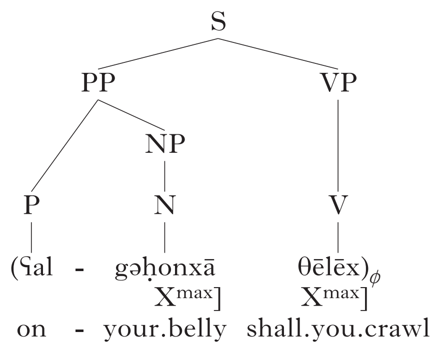
Prosodic weight also plays an important role in the derivation of Tiberian Hebrew. Typically, a Biblical Hebrew phonological phrase consists of two prosodic words. Therefore, a syntactic unit that would normally be divided into two phrases forms a single phrase when it consists of only two words. In this connection, compare (16a) and (16b).

(16) *Effects of weight on phrasing*

a. *Gen.* 3.16: *Two phrases*



b. *Gen.* 3.14: *One phrase*



(16a) has a phrase boundary at the right edge of the PP, as we expect. However, (16b) unexpectedly has only one phonological phrase. In (16b), the preposition has been cliticised to its object. Since only two words remain, they are combined into one phrase, annulling the phrase boundary that otherwise occurs in this position.

Conversely, syntactic units which are normally kept together if they consist of two words are broken up if they contain more than two words. An example of this can be found in (14a) above, where the verb is separated from its branching complement.

Another example of this weight balancing occurs in (17). The typical phrasing of a verb with a preceding lexical subject is shown in (17a), where the subject forms one phrase, and the verb and its complement form another phrase. However, when the verbal complement contains two or more words, the verb must form a separate phrase; in certain circumstances, it may be phrased with a one-word subject, as in (17b):

(17) *Phrasing of subject–verb–complement sequences*

a. (wəlōṭ) (yōšēv bəšaʕar-səðōm)
and.Lot sat in.the gate (of)-Sodom (*Gen.* 19.1)

b. (wəʕefrōn yōšēv) (bəθōx bənē-ḥēθ)
and.Ephron sat in.the.midst (of) the.children (of)-Heth (*Gen.* 23.10)

A key observation for our present purposes is the following: in certain contexts, the Hebrew phonological phrase consists of only one word, even when two words are available. Consider (14b): given the above principles, we would have expected the whole sentence to form a single phonological phrase, like the parallel (16b). The difference is that the phrase in (14b) is at the right edge of an intonational phrase. In final position of an intonational phrase, the principle in (18) applies:

(18) *Division Rule*

A two-word phonological phrase in a prominent position of a prosodic tree (the head of an intonational phrase) is divided if one of its words is long.

A number of terms in the Division Rule need to be further explicated. The rule mentions a phonological phrase in a prominent prosodic position. In terms of the Tiberian system of notation, such a phrase is identified as being the rightmost phrase under a high-level accent, interpreted by Dresher (1994) as designating an intonational phrase. That the rightmost phrase in an intonational phrase is in fact the most prominent one is attested to by various phonological strengthening phenomena, such as vowel lengthening and heightened stress, which characteristically occur here. In other terms, we can say that the rightmost phonological phrase is the head of its intonational phrase, and the other phrases are dependents.

The Division Rule also refers to the notion of LONG WORD. What is meant here is a rather specific definition of what counts as long, given in (19):

(19) *Long word*

A long word is one which has at least two vowels before the stressed vowel, not counting reduced vowels (i.e. schwa or ultra-short vowels considered to be variants of schwa); or else contains a vowel marked with secondary stress.

Dresher (1981a, b, 1982) has argued that the notions of 'long' and 'short' word can be best understood in metrical terms. A word marked with secondary stress has two metrical feet. Such marks are often assigned to a non-final superheavy syllable, which is sufficient to support a metrical foot. Apart from such cases, and syllables with primary stress, which can also support a foot by themselves, it can be shown that a Biblical Hebrew foot ordinarily requires two full syllables. Therefore, the relevant notion of long word is a word which consists of at least two feet.

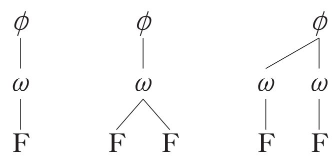
(20) *Metrical interpretation of 'long word'* (Dresher 1981a, b, 1982)
A long word has at least two left-dominant feet, where a foot is constituted by:
a. the syllable with main stress and any following syllable;
b. every two full syllables in an alternating pattern before a stress.

Thus, the prosodic word [wəʔel-ʕāfā́r] in (14a) is long, according to the above definitions; since the following word is final in the intonational phrase, it is not phrased together with the long word, but each word forms a separate phonological phrase.

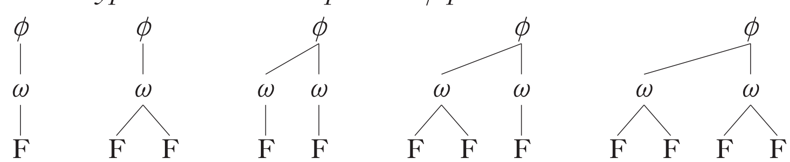
2.2 Another type of HDA: visibility

Let us consider now how we might account for phenomena like the Division Rule. We observe first that it involves an asymmetry between a head phonological phrase and dependent phonological phrases in terms of how much complexity each allows. This asymmetry, however, goes exactly the wrong way: the head phonological phrase allows for less complexity than dependent phrases do, as it made clear by (21):

(21) *The Division Rule as an anti-HDA*
a. *Word types allowed in a head ϕ-phrase*

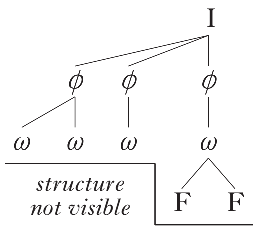


b. *Word types allowed in a dependent ϕ-phrase*



Although (21) accurately characterises the facts, it does not explain why such an asymmetry should exist. We propose that there is another way of looking at the Division Rule which allows us to better understand its rationale. As a first step, we observe that in prominent prosodic positions, it is as if the branching at the foot level counts as if it were branching at the word level. Thus, a word with two feet counts as if it is two words, exhausting its phonological phrase. A formal characterisation of just such a situation is given by Vergnaud & Halle (1979), who propose a notion of incorporation: in prominent positions, the immediately lower level can be incorporated into a prosodic tree. This situation can be schematically diagrammed as in (22):

(22) *HDA under the I-phrase* (*Tiberian Hebrew*)



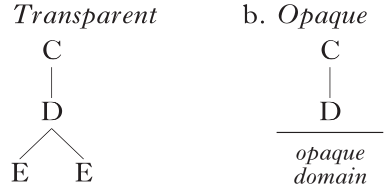
The line represents the bottom level of structure that must be analysed in the formation of each phrase. Dependent phrases look as far as the word level; the head phrase looks further, to the foot level. Thus, a word with two feet counts as branching in terms of the phonological phrase level. The head phrase, then, cannot be composed of two prosodic words if one of these is bipodal.

Notice that we still have an asymmetry between a head and its dependents, but in this case it is not one of complexity, but of visibility: head phrases are sensitive to further levels of structure. In dependent positions, all words are treated the same, i.e. as one unit, irrespective of their internal make-up. In head positions, words are not equal. Unlike the characterisation in terms of complexity, this HDA goes in the expected direction: heads require a deeper, more articulated analysis of the structure than dependents.

Tiberian Hebrew thus involves a second type of HDA. In general terms, this type of HDA involves the degree to which a phonological process which applies to a node C has access to the internal structure of its daughters. In (23), we will say that C is (relatively) transparent if it allows access to nodes of type E, and opaque if nothing can be 'seen' below the level of D:

(23) *Visibility HDA*

a. *Transparent* b. *Opaque*



Our claim is that where there is a difference between a head and its dependents with respect to visibility, constituent units of the head are visible to a greater depth than the corresponding constituents of the dependents. We will call this type of asymmetry a VISIBILITY HDA (HDA-V).

2.3 Ligurian Italian

This kind of asymmetry is quite common. An HDA-V similar to that of Tiberian Hebrew is found in Ligurian Italian. According to Ghini (1993a, b), phonological phrases in Italian are formed as in (24):

(24) *Formation of Italian ϕ-phrases* (Ghini 1993a, b)

a. *ϕ-domain*: The domain of ϕ-formation is delimited by right-edge X^{max} boundaries.

b. *ϕ-formation*: Phonological words included in the ϕ-domain are distributed according to the following principles:

(i) Average weight: each ϕ ideally contains two phonological words.

(ii) Symmetry: symmetric prosodic units are preferred.

(iii) Increasing units: in asymmetric structures, prosodic units on the right are heavier than prosodic units on the left.

Words within the domain of a maximal projection are organised into phrases so as to achieve symmetrical two-word phrases if possible, as in (25a), which has two two-word phrases; in the case of an odd number of words, the extra word is preferably assigned to the rightmost phrase, as in (25b), not (25c). In these examples, forms joined by hyphens comprise a single prosodic word:

(25) *Italian phrasing* (based on Ghini 1993a, b)

a. (Ho-studiàto tútto) (il-lìbro di-tésto)
'I have studied the whole textbook.'

b. (Ho-studiáto) (tùtto il-líbro)
'I have studied the whole book.'

c. *(Ho-studiàto tútto) (il-líbro)

When a word in phrase-final position is long, however, the expected phrasing (26a) is not found; instead we find (26b):

(26) *Italian phrasing with long words in head position of ϕ*

a. *(Ho-studiáto) (tùtta la-rìvoluzióne)
'I have studied the whole revolution.'

b. (Ho-studiàto tútta) (la-rìvoluzióne)

If the long word *la-rivoluzióne* is treated as an ordinary word, the pattern in (26) is hard to explain. However, if a long word (presumably a word with at least two feet) is counted as two words, then (26b) would exemplify the same balanced pattern as (25a), i.e. (ω ω)(ω ω). The phrasing in (26a) is ruled out as being too unbalanced, as if it were (ω)(ω ω ω).

Now, it turns out that long words are not always treated as being equivalent to two words for purposes of phrasing: when a long word is not phrase-final, it is counted as one word, as demonstrated in (27):

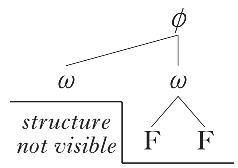
(27) *Italian phrasing with long words in dependent position of* ϕ

a. (Ho-studiáto) (la-rivoluziòne francése)
   'I have studied the French revolution.'

b. ?(Ho-studiáto) (la-rìvoluzióne) (francése)

If *la-rivoluzióne* were counted as two words in (27), then (27a) would exemplify the usually avoided pattern (ω)(ω ω ω), parallel to (26a), and we might expect (27b) to be preferred. But this is not the case: (27a) is treated just like (25b).[12] These facts can be explained if the internal composition of a long word is relevant to phrasing only in head position of a phonological phrase:

(28) *HDA under the* ϕ*-phrase* (*Ligurian Italian*)



The difference between Biblical Hebrew and Italian is that in Biblical Hebrew the length of a word is projected only in the head phonological phrase of an I; the length of either word in this phrase is relevant. Thus, this is an HDA at the level of the intonational phrase, affecting the head of I. In Italian, the length of the head word of any phonological phrase is relevant, but not the length of the dependent word. Hence, this is an HDA at the level of the phonological phrase, involving the head of ϕ. In terms of our earlier distinction between local and non-local HDAs, the Tiberian Hebrew case is non-local, whereas the Italian HDA is local. In the latter case, the ϕ-level HDA involves the internal structure of the immediate daughter (ω) of ϕ, whereas the Tiberian Hebrew I-level HDA involves the internal structure of a non-immediate daughter (ω) of I. Thus, the local *vs.* non-local distinction is relevant for visibility HDAs as well as for complexity HDAs.

3 Visibility *vs.* complexity

It is important to distinguish between visibility and complexity HDAs. In the former case shown in (23), the asymmetry has only to do with accessibility: the internal structure of D in (23b) may be equal in complexity to that of (23a), but since it is not accessible to the higher levels

[12] We would like to thank Mirco Ghini for supplying these examples.

of structure, any complexity it may have does not count at the higher level.[13] In the non-local complexity HDA shown in (3), a remote daughter of the head (3a) is in fact more complex than a remote daughter of a dependent (3b). In such cases, heads may require a certain complexity 'lower down' in the hierarchy, whereas its dependents do not.

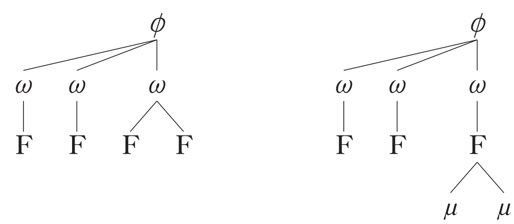
So far, we have seen cases of complexity HDAs at the level of the foot, and visibility HDAs at the level of the phrase and higher. Is it, then, the case that the two types of HDA are in complementary distribution, with one occurring at lower levels of the prosodic hierarchy and the other at higher levels? The answer appears to be no, for we can also find examples of complexity HDAs at the higher levels, as we will show in the next section.

3.1 Phrase-level complexity HDAs: head *vs.* dependent words

Some complexity HDAs at the phonological phrase level are schematically illustrated in (29). Example (29a) illustrates a language where the word that occupies the head of phrase, say the rightmost word in a phrase, may have a degree of complexity greater than that permitted to phrase-internal words. For example, the head word, but not the dependent words, might be required, as in (29a), to consist of at least two feet; or, in the non-local version of this HDA in (29b), the head word might be required to have at least a branching foot, i.e. a foot dominating two moras, a degree of structural complexity not required of (or allowed to) dependents:

(29) *Complexity HDAs at the word level* (hypothetical)

a. *Local HDA* b. *Non-local HDA*

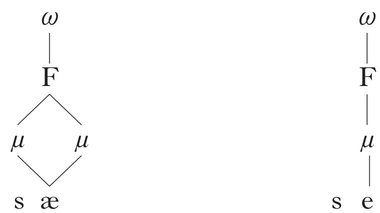


Example (29b) in fact illustrates the typical case of minimal word requirements found in many languages, whereby a word must have a certain complexity (McCarthy & Prince 1990). Such requirements are typically imposed on content words, i.e. words that can be the heads of maximal projections; other words, such as function words, i.e. words belonging to certain closed classes, are often excluded from these requirements. For example, a minimal word in Old English must have at least two moras, a requirement that is imposed on major class words such as nouns, e.g. *scip* 'ship', *sæ:* 'sea', **sci*, **sæ*, verbs and adjectives (30a); however,

[13] Compare the discussion of constraints on depth in connection with Chomsky's Strict Locality Principle, called the Head Constraint by Hoekstra *et al.* (1980: 37–38).

minor class words (*we* 'we', *se* 'that', *hwa* 'who?') may escape it, so that the words in (30b) are permissible:

(30) *Lexicalised HDAs at the word level* (*Old English*)
a. *Major class words* b. *Minor class words*



We propose that this type of case represents the lexicalisation of a ϕ-level HDA: major class words, which are all potential heads, must be complex all the time, not just when they actually are in head position. Minor class words, which are never heads, may be exempt from minimality.

We could imagine a situation in which dependent words would be required to be non-branching, but we do not expect to find this kind of asymmetry for major class (content) words, because such a situation would require pervasive allomorphy, whereby all major class words would come in two forms: a longer form for head position, and a shorter form for other positions. Moreover, the allomorphy would have to be productive, applying to new and morphologically complex words, if the system is to maintain itself.

Hence, such a situation would be very costly if it were applied to major class items. The same, however, is not true of closed classes, and we do find examples where minor class words come in two forms, strong and weak, whose distribution has a prosodic basis. An example is Modern English, where modals, possessives, determiners and so on come in strong and weak varieties. The weak varieties occur in phrase-medial or even cliticised positions, and often fall below the minimal weight requirement for regular English words: e.g. *n't* for *not*, *him* reduced to syllabic [m̩], etc.[14]

To conclude this example, then, we consider the existence of asymmetric minimality requirements to be lexicalised versions of complexity HDAs. Closed class items have forms which are not subject to minimality because these either do not occur as prosodic heads or, if they do, the lexicon can afford to have two forms, a weak and a strong form.

We would expect to also find corresponding complexity HDAs at the level of the intonational phrase. In such an HDA the head phrase of an intonational phrase (usually the final ϕ in an I) would optionally or obligatorily have more structure than dependent phrases. Such an HDA could be local, referring to the words making up a phrase, or non-local, involving some distinction at another level of structure. Yip (1989)

[14] These alternations are discussed by Selkirk (1972).

observes that extra-complex contour tones are found only in domain-final position in Chinese dialects, as well as in the African tone language Mende; utterance-internally they are simplified in various ways.[15] Various other phenomena are known to mark intonation peaks, and these phenomena supply a pool of possible cases of this type.

It is apparent, then, that complexity and visibility HDAs are not in complementary distribution, but can both occur at the same levels of the prosodic hierarchy.

3.2 An apparent HDA paradox

The two types of HDA both attribute properties to a head that are not found in dependents. In the case of complexity, a head is allowed to be more complex than a dependent, but not *vice versa*; in the case of visibility, a head may require a more articulated analysis of structure than a dependent does, but not *vice versa*. Each of these HDAs embodies an empirical claim that appears to reflect a common situation. Individually, they are unproblematic (though that is not to say they are self-evidently true); but taken together, they raise an interesting paradox.

We have already had a glimpse of this HDA paradox in our discussion of the Hebrew Division Rule: though it is a perfect example of a HDA-V (22), it can also be viewed as an HDA-C (21) that goes in the wrong direction, i.e as an 'anti-HDA-C'. The same is true of Italian. We observed that Italian phrases ideally consist of two words (at a normal rate of speech), except when the head word is long, in which case this long word seems to exhaust the space that phrases have available to them. Whether dependent words are long or short does not seem to matter. Our interpretation of this situation is that long words have at least two feet, and that the head of a phrase is transparent in some sense, allowing the branching feet to be visible to phrase well-formedness constraints, whereas the dependent word is opaque. Implicit in this view is the claim that the phenomenon attested in Italian does not involve a complexity asymmetry between the daughter words in a phrase. Attributing the asymmetry in Italian phrases to visibility is in keeping with the intuition that, if such an asymmetry exists, it is expected that heads will make their internal structure visible, whereas dependents will not.[16]

However, one might be tempted to incorrectly characterise this asymmetry in another way: in Italian, head words must be short if a dependent is present (i.e. in a two-word phrase), whereas a dependent word can be long or short. In other words, it might seem that there is a complexity limit on heads which does not exist for dependents:

[15] Yip does not indicate whether the relevant domain in these cases is the utterance or the intonational phrase; we are assuming here that it is the latter. Such phenomena may of course occur at various levels of the prosodic hierarchy.

[16] This intuition, in fact, forms the backbone of X-bar theory, which is entirely based on the idea of non-local visibility in the form of the projection/percolation relation between mother nodes and their heads.

(31) *Italian phrasing interpreted as a complexity HDA*
  a. The head of a phrase must be short if a dependent is present.
  b. Dependents in a phrase may be long or short.

The characterisation in (31) is an 'anti-HDA-C': an asymmetry in complexity which goes exactly in the wrong direction. More generally, we are faced with the apparent paradox that visibility and complexity lead to (almost) opposite predictions in most cases.

In its most general form, then, the HDA paradox can be stated as follows: by virtue of being a head, an element may be expected to have greater complexity than a dependent (HDA-C); but at the same time, a dependent may be expected to have greater complexity than a head because it is so lacking in prominence that any complexity it may have is disregarded, so that it may escape from complexity constraints that apply to the head (HDA-V). There is a fine line between a unit which is capable of great complexity in a certain dimension and one which in a sense has no complexity at all, being insensitive to the relevant dimension.

It is, nevertheless, one thing to understand why complexity and visibility go in opposite directions, and another thing to have a principled way of distinguishing between them. For if we are free to assign any case to one or the other category, the claim that each embodies becomes quite empty.

3.3 Visibility and mapping relations

We believe that there is a fundamental difference between visibility and complexity asymmetries which can be established independently, so that we are *not* in fact free to choose whether we appeal to one or the other, depending on the way the asymmetry goes.

Let us consider again the Division Rule in Biblical Hebrew, and try to see what might underlie the visibility HDA that exists there. Janis (1987) has referred to such rules as 'pacing rules', because they regulate the speed at which words are said. When a word is said more slowly, it takes more time. If we assume that there is a rough isochrony of phonological phrases in the Biblical Hebrew reading tradition recorded in the texts, it would follow that a phrase can include more words that are said quickly, but fewer words that are said slowly.

The introduction of time as an independent unit leads us to a fundamental re-evaluation of the syntax–prosody mapping. In earlier formulations of this mapping (Nespor & Vogel 1986, Hale & Selkirk 1987), the only constraints on prosodic structure (at least at the level of the phonological phrase) are syntactic:

(32) *Phonological phrase formation* (Nespor & Vogel 1986: 168)
  a. *ϕ-domain*: The domain of ϕ consists of a clitic group (*c*) which contains a lexical head (X) and all *c*'s on its non-recursive side up to the *c* that contains another head outside of the maximal projection of X.

b. *ϕ construction*: Join into an n-ary branching ϕ all *c*'s included in a string delimited by the definition of the domain of ϕ.
c. *ϕ relative prominence*: In languages whose syntactic trees are right branching, the rightmost node of ϕ is labelled *s*; in languages whose syntactic trees are left branching, the leftmost node of ϕ is labelled *s*. All sisters nodes of s are labelled *w*.

(33) *End-based mapping of phonological phrases* (Hale & Selkirk 1987: 152)
a. *The Designated Category Parameter*: For each level P_i of the prosodic hierarchy there is a single designated category DC_i in the syntactic structure with respect to which phonological representation at level P_i is defined.
b. *The End Parameter*: Only one end (Right or Left) of the designated category DC_i is relevant in the formation of a prosodic constituent P_i; a P_i extends from one instance of the appropriate end of DC_i to the next.

In subsequent work, prosodic conditions take on more prominence: compare Ghini's formulation given above in (24), and the complex weight effects in Tiberian Hebrew. Now it appears that some conditions which have been stated in purely prosodic terms also involve the notion of timing units.

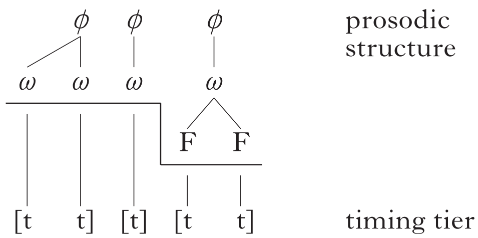
The fact that one can go a long way in stating prosodic constraints without reference to an independent timing unit is due, on our view, to the fact that timing units are typically coordinated with units of prosody. For example, if in a certain language a phonological phrase optimally takes up an amount of time equivalent to two words, one can just say that a phonological phrase optimally has two words, without referring to time.

In more complex situations, however, the equivalence between timing and specific units of prosody may break down. In particular, we propose the hypothesis in (34):

(34) *Relation of timing to prosody*
Where a level of the prosodic hierarchy, P_i, is subject to timing constraints, the relevant timing units are calibrated in terms of the prosodic unit next lower down, P_{i-1}. Deviations from this calibration may occur at non-standard speech rates, or in prominent or subordinate positions in the utterance.

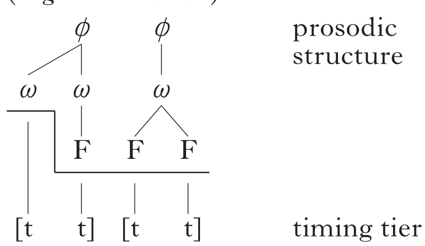
Consider how (34) applies to the Hebrew and Italian cases. In Hebrew, we have assumed a rough isochrony of phonological phrases. This isochrony is usually measured in terms of words: the standard length of a phonological phrase is two words. However, as we have seen, in the head phonological phrase, the level at which weight is computed shifts from the word to the foot: the head phonological phrase aims at a length of two feet. This shift corresponds to a slowing down of the speed of the words in the head phrase, if the overall time for a phonological phrase is maintained. It follows that timing cannot be computed in terms of words or feet, but must be computed on an independent plane onto which words and feet are mapped, along the lines of (35):

(35) *HDA-V as a mapping asymmetry between two planes (Biblical Hebrew)*



In the Italian example, the units that are mapped onto the timing plane are words in dependent position, but daughters of words, i.e. feet, in head position. *Vis-à-vis* the mapping, then, it is as if heads are 'magnified'. This is the nature of visibility asymmetries: heads may be magnified in this way relative to dependents, but dependents may not be magnified relative to heads:

(36) *HDA-V as a mapping asymmetry between two planes (Ligurian Italian)*



Our picture of the constraints on prosodic structure is now as follows:

(37) *Constraints on prosodic structure*

$\boxed{\text{syntax}} \longleftrightarrow \boxed{\text{prosodic structure}} \longleftrightarrow \boxed{\text{timing}}$

We propose that the fundamental distinction between the two types of HDA is this: complexity asymmetries involve a comparison of daughters of a constituent *within a single plane*; visibility asymmetries arise in the *mapping from one plane to another*. As long as we stay within a single plane, only HDA-Cs are permissible, and dependents cannot be more complex than heads.

For example, in the word-minimality example given above in (30), the relevant constraint is that a head word must include a complete foot. We are not free to ignore foot structure as a way of evading this constraint. The Italian case, by contrast, involves a relation between two phonological planes: one is a plane which encodes prosodic hierarchical organisation, and the other involves timing. We will refer to the latter plane as the *timing tier*, or *grid plane*, having in mind a notion of grid that goes back to Liberman & Prince (1977), Prince (1983) and Hayes (1984). This plane represents rhythmical beats that are timed so that they are separated by roughly even time intervals.

We maintain that a consistent HDA syndrome exists for each type of asymmetry. In complexity asymmetries, the syndrome is that dependents may not be more complex then heads (either syntagmatically or paradigmatically), whereas in visibility asymmetries mapping may not refer to daughters of dependents unless they also refer to daughters of heads.

In the next section, we consider another type of situation where heads may be less complex than dependents. Unlike the cases discussed above, where the violation of the HDA principle is only apparent, being due to visibility asymmetries, the next class of cases result in real violations. Such violations do not occur arbitrarily, however, but are due to another principle which treats heads as special.

4 Heads first!

In this section we examine a number of cases that involve word-level asymmetries. The asymmetries lie in the fact that head and dependent feet differ in quantity-sensitivity. First, we address the case in which head feet are QS while dependent feet are QI. We argue that this is a visibility asymmetry of the expected kind, where the head foot (which bears main stress) is sensitive to a deeper level of analysis than the dependent feet. Then we turn our attention to cases that work the other way around. Such cases are unexpected, given our remarks to here, since they imply an anti-HDA-V. We will argue that such cases involve the main-stress first syndrome (van der Hulst 1984, 1996c), or what other have called top-down parsing (Hayes 1995).[17]

4.1 Visibility effects

There is a phenomenon found in a number of languages whereby main stress appears to be sensitive to quantity (QS), but secondary stress is not

[17] It may be useful to ask whether we also have HDA-Cs at this level. An asymmetry of this kind would arise if, for example, head feet must be branching, while dependent feet can be branching or non-branching. We are aware of one situation which closely fits this description. In early analyses of Dutch stress (van der Hulst 1984), the generalisation with respect to which the foot heads the word is that the rightmost foot is strong unless it is non-branching. In this case the penultimate foot would be the head, whether it branches or not. More recently, phonologists have appealed to foot extrametricality to achieve the same effect. We do not wish to investigate this issue in further detail here, but it seems clear that the specific requirement that the head foot must satisfy potentially qualifies as a local HDA-C at the word level; cf. Kager (1989) and Trommelen & Zonneveld (to appear) for an overview of the literature.

(QI). Examples are Seneca (Stowell 1979, Halle & Vergnaud 1987) and Maung (Capell & Hinch 1970):

(38) *Main stress–secondary stress asymmetry*

a. *Seneca*: Main stress falls on the last non-final even-numbered syllable that is either closed or immediately followed by a closed non-final syllable; secondary stress alternates in an iambic pattern preceding the main stress.

b. *Maung*: Main stress falls on the first two syllables of words of up to three syllables (equally stressed), and on the penultimate of longer words, except that a preceding closed syllable tends to take primary stress; secondary stresses fall on alternate syllables preceding main stress.

In these languages, main stress depends on syllable weight, while secondary stress then alternates back from the main stress in a quantity-insensitive fashion. Halle & Vergnaud (1987) find this situation to be so common that they give it a special name, the Alternator.[18]

A variant of this asymmetry in quantity-sensitivity is found in languages like Spanish (Roca 1986, Halle *et al*. 1991), Italian (Vogel & Scalise 1982) and Chamorro (Chung 1983, Halle & Vergnaud 1987). In these languages, main stress falls on one of the last three syllables of the word in an unpredictable fashion, and secondary stress alternates evenly on preceding syllables. To account for main stress in such languages we may suppose that certain syllables may be exceptionally marked for accent. Such cases are like quantity-sensitivity in that construction of metrical structure depends on some further property of syllables beyond just their linear sequence.

These facts can be readily represented in terms of current versions of metrical theory, but they remain to be accounted for. By now the explanation for this tendency should be evident: it is a visibility HDA at the level of the word. The foot bearing main stress in a word is a head, whereas feet controlling secondary stress are subordinate to it. Only the head foot is built on lower-level projections from syllable structure in which the distinction between light and heavy syllables is maintained. The non-head feet have no access to this distinction, and are rather built on projections which do not differentiate lower-level syllable structures.

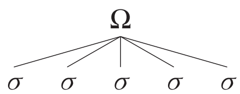
Since these cases involve visibility asymmetries, it follows from our hypothesis that two planes must be involved. We will now develop a view on the relation between the prosodic plane and what we will call the WORD PLANE which entails that feet are not built on syllables but rather on metrical positions which correspond to either syllable rhymes or to rhyme daughters.[19]

[18] In an interesting analysis, Pater (1994) shows that English secondary stress is sensitive to weight, but that the weight of heavy syllables can be suppressed when higher-ranked constraints supervene. Weight suppression, however, arguably never takes place in syllables in the main stress foot.

[19] Our position here is close to that found in Halle & Vergnaud (1987).

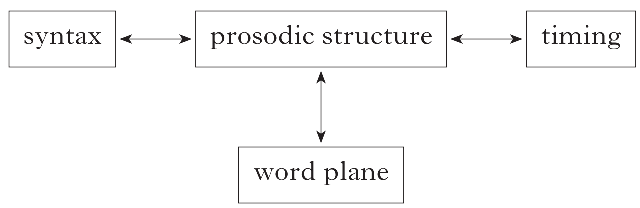
We assume that the syllabic organisation of segments is represented in a separate word plane. This plane is distinct from the prosodic plane.[20] The word plane consists of a sequence of syllables that forms a word:

(39) *Word plane*



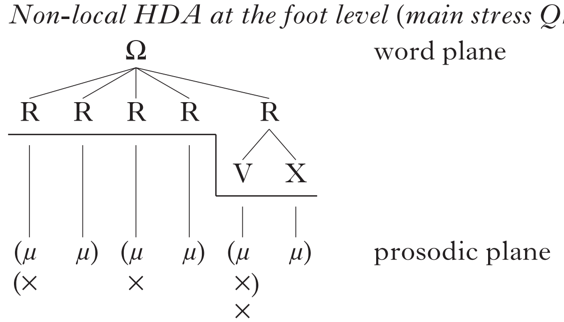
There is a mapping between the word plane and prosodic structure.[21] Thus, our diagram (37) above should be augmented as in (40):

(40) *The word plane and prosodic structure*



The terminal elements of prosodic structure, we assume, are line 0 units that do the work that in other theories is attributed to moras (μ). On this interpretation, feet dominate moras directly, syllables being in another plane. Between the word plane and the prosodic plane a mapping relation exists whereby line 0 units correspond either to rhymes (QI) or to daughters of rhymes (QS). It follows, then, that if a QS/QI asymmetry exists, we expect to find that dependent feet are built on line 0 units that correspond to complete rhymes, whereas head feet are built on daughters of the rhyme, as shown in (41):

(41) *Non-local HDA at the foot level* (*main stress QS, alternator QI*)

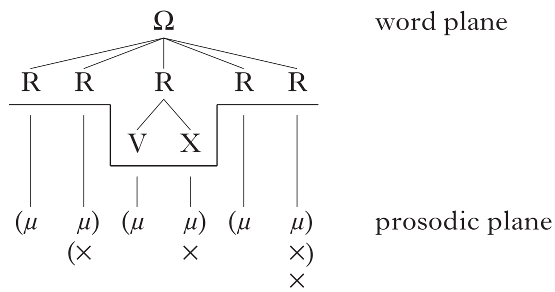


[20] On the distinction between word plane and prosodic plane see Anderson & Ewen (1987: 122ff), van der Hulst (to appear a) and van der Hulst & Rowicka (1997).

[21] The word plane is the locus for phonosyntactic constraints. As such, it may contain 'foot-like' relations between syllables or rhymes as envisaged in government phonology; see van der Hulst & Rowicka (1997) for discussion. We represent a word in the word plane as Ω, to distinguish it from the prosodic word, ω.

The asymmetry is as expected: mapping from the word plane to the prosodic plane may not refer to daughters of dependents unless they also refer to daughters of heads. Hence, the contrary situation, where secondary stress is sensitive to quantity while main stress is not, is ruled out, since it would require non-heads to be sensitive to a distinction that is not available to the head:

(42) *An ill-formed mapping*



4.2 The main stress first syndrome

The literature on stress systems reports a number of cases in which main stress is located on a fixed (near-)peripheral syllable. Two well-known cases are Finnish and Old English.

Finnish has fixed initial primary stress, i.e. syllable weight has no influence whatsoever on its location. The pattern of secondary stresses, however, reveals weight sensitivity. The first secondary stress falls on the third syllable (43a–c), except when it is light and the fourth syllable is heavy (43d). In this case the first secondary stress falls on the fourth syllable:

(43) *Finnish stress patterns*

a. $\acute{\sigma}$ σ $\grave{\bar{\sigma}}$ $\breve{\sigma}$
b. $\acute{\sigma}$ σ $\grave{\breve{\sigma}}$ $\breve{\sigma}$
c. $\acute{\sigma}$ σ $\grave{\bar{\sigma}}$ $\bar{\sigma}$
d. $\acute{\sigma}$ σ $\breve{\sigma}$ $\grave{\bar{\sigma}}$

$\bar{\sigma}$ = heavy syllable
$\breve{\sigma}$ = light syllable
σ = any syllable

It would seem that the head foot is quantity-insensitive whereas the dependent feet show quantity-sensitivity of the non-mismatch type (Grijzenhout 1992).

The Old English stress system has been the subject of a great many studies (see Dresher & Lahiri 1991 for an example). The primary accent is fixed on the first syllable. To account for a number of phonological processes, however, it seems clear that the weight of syllables beyond the first syllable matters for how foot structure is assigned.

Examples of this kind can be multiplied; see van der Hulst (1996c) for further cases. Hayes (1995: 116) proposes that such cases involve 'top-down parsing'. In recent approaches within Optimality Theory (Prince & Smolensky 1993) it is argued that top-down effects prove the non-derivational character of phonology. The effect arises when constraints bearing on the word level (i.e. on the location of main stress) outrank those bearing on the foot level (i.e. on secondary stresses). Van der Hulst (1984, 1996c) adopts the view that in such cases primary stress is assigned at an earlier grammatical level than rhythmic structure; the former is lexical, whereas the latter is postlexical. It is the latter view that we wish to integrate into the approach that we advocate here.

If assigned lexically, primary accent forms part of the lexical structure, i.e. it becomes part of the word plane. Rhythmic structure, on the other hand, forms part of the (postlexical) prosodic plane. Some information from the lexical plane is projected onto the postlexical plane; in languages like Finnish and Old English, the head of the lexical prosodic structure (the main stressed syllable) is projected as a head in the prosodic structure. Rhythmic foot formation must then respect the location of the lexical heads.

In §4.1, we observed that there are some cases where main stress is QS and secondary stresses are QI, an expected type of visibility HDA. The cases discussed in this section are thus problematic, for they go in the opposite direction (main stress QI; secondary stress QS), and hence represent an anti-HDA. However, we have argued that in these cases, the principles that determine primary stress are independent of and dominant over principles that determine rhythmic stresses. Thus, main stress and secondary stress feet are not derived from the same projection; hence, these asymmetries have no bearing on the nature of HDAs.

5 Further aspects of HDAs

5.1 HDA in the phonology of two-handed signs

HDAs are manifested quite clearly in the organisation of signs, and here we will explore one way in which such asymmetries are realised.[22]

Many signs make use of both hands rather than just one. To refer to the hands, we will follow Padden & Perlmutter (1987), who use the terms 'strong hand' and 'weak hand' rather than 'preference hand' and 'non-preference hand'. The preference hand (i.e. the left hand of left-handed signers and the right hand of right-handed signers) will normally be the strong hand, but this is not a necessity. Van der Hulst (1993, 1995, 1996b) refers to signs in which the weak hand is a place as 'unbalanced', whereas the other type is referred to as 'balanced'.

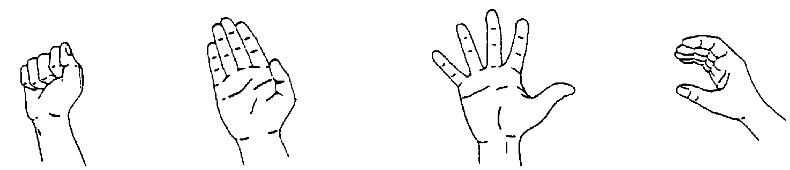
Since Stokoe (1960), researchers have adopted an analysis of signs in terms of a number of 'class nodes', such as Place, Handshape, Location and Movement. For the purpose of this discussion, we can say that in

[22] More extensive discussions of two-handed signs can be found in van der Hulst (1996b), Sandler (1995) and van der Hulst & Sandler (1994).

balanced signs both hands have equal specification for all class nodes. In unbalanced signs, the strong hand is fully specified for the class nodes, but this is not the case for the weak hand. It is a striking fact that if the handshapes differ, the weak hand has a choice from a very limited set (at least among the most frequent cases):

(44) *Weak handshapes in ASL unbalanced signs*

Fist Flat Flat-spread fingers C-hand



The handshapes that may be selected are unmarked. According to Battison (1974: 6–7) (and cf. Sandler 1995), these handshapes:

- are maximally distinct both in articulatory and perceptual terms;
- have a high frequency of occurrence in a wide variety of contexts;
- are found in all other sign languages studied to date;
- are among the first handshapes acquired by children;
- are the outcome of substitution errors;
- are capable of contacting other body parts in a greater variety of ways.[23]

In the current literature, there are differing views as to how to represent two-handed signs (see van der Hulst & Sandler 1994). One question which arises is how the weak hand is represented in balanced and unbalanced signs. The most conservative position is that which somehow or other represents the weak hand separately in all two-handed cases. Most recently, this position has been taken in Brentari (1990) and van der Hulst (1996b). Sandler (1994) proposes to represent only the weak hand in balanced signs.

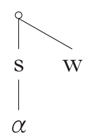
Van der Hulst argues that the asymmetry between strong and weak handshapes fits the general HDA pattern rather well. If we assume that the two hands enter into a head–dependent relation in which the strong hand is the head, the limitation on weak handshapes is entirely in line with the general asymmetry between heads and dependents. In balanced signs, the weak hand is specified as a copy of the strong hand; we represent this here in terms of spreading (45a). The class of weak handshapes that may occur in unbalanced signs forms a subset of the class of strong handshapes. This phenomenon is in fact very reminiscent of the widespread asymmetry between strong and weak syllables in foot structure. In such cases, as we have seen above, the set of vowels which may appear in weak syllables is a proper subset of the set that occurs in strong positions (45b). Moreover,

[23] See Sandler (1995) for a discussion of the markedness issue and a formal characterisation of unmarked shapes.

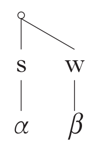
the subset occurring in the weak position counts as the unmarked set. In balanced signs, we find an even more dramatic reduction for the dependent hand: all contrast has been removed.[24] Clearly, if the weak hand-as-place is represented under the place node, we have no formal basis for explaining this asymmetry.

(45) *Balanced and unbalanced handshapes*

a. *Balanced two-handed sign*



b. *Unbalanced sign*



5.2 Implications for acquisition

In his article on degree-0 learnability, Lightfoot (1989) explores the possibility that the syntax of a language is learnable only from main clauses, without recourse to data from embedded clauses. He notes that the proposal is attractive on several grounds. From the point of view of acquisition, this hypothesis, building on work of Wexler & Culicover (1980), claims that grammars should be learnable on the basis of simple and ubiquitous data which every child has access to. He then goes on to observe that an acquisition strategy of this kind could also account for a generalisation noted by Ross (1973) and Emonds (1976), to the effect that many transformations can be found which are limited to root clauses, whereas there are arguably no transformations which are limited to embedded clauses. Lightfoot points out that this asymmetry follows if learning is based primarily on main clauses, because then no operation would be expected in a subordinate clause which did not also occur in a main clause.

Lightfoot is not able to confine learning to degree-0, however, and concludes that a certain amount of embedded clause information is required also, notably data involving complementiser properties. He concludes that what is needed is 'degree-0 and a bit'. In his commentary on this article, Rizzi (1989) points out the rather arbitrary nature of 'degree-0 and a bit', and proposes that the extra bit appears to be limited to information about syntactic heads. He proposes that children pay attention to heads, and that the grammar is to a large extent learnable from the properties of heads.

Among his strategies of acquisition, Slobin (1973) proposed maxims such as 'pay attention to stressed syllables', 'pay attention to the beginnings (or ends) of words' and so on. The question arises as to what it is about these things that makes them part of a class to which learners should pay special attention. Our proposal is that they are all heads at

[24] In a somewhat similar spirit, Brentari (1990) represents the weak hand as a kind of 'coda', deriving the limited number of contrasts as a 'coda weakening effect'.

some level of the grammar, and that many of these maxims can be subsumed under a more general one: 'pay attention to heads!'.

As in the Lightfoot–Rizzi hypothesis for syntax, we locate the origin of HDAs in the acquisition process. We assume that learners begin with relatively impoverished representations, and move to more richly articulated representations under the pressure of data (cf. Dresher to appear). Moreover, the strategy of 'pay attention to heads' implies that heads will be expanded before dependents. In many cases, the dependents catch up; but when they do not, the result is an HDA.[25]

We have not discussed the implications these facts might have for the developmental course of acquisition; however, if this investigation is on the right track, we expect that the empirical study of stages of acquisition will shed a great deal of light on head–dependent relations.

6 Conclusion

In this article we have developed a typology of head–dependent asymmetries (HDAs). In one type, complexity HDAs, heads tolerate a greater degree of complexity than dependents. This has in fact been quite generally observed. We have shown that a careful examination of the relevant examples requires us to distinguish between local and non-local HDAs, and particularly between paradigmatic and syntagmatic HDAs. We have also pointed to the distinction between α-α (where head and dependents are constituents of the same type) and α-β (heterogeneous) HDAs.

Our main interest has been to understand phenomena that appear to involve HDAs of the wrong kind, i.e. cases in which heads appear to be less complex or less sensitive than dependents. We have argued that an important class of such cases involves an HDA asymmetry that relates to the notion of visibility. Such cases, as we have shown, arise when information is projected from one plane to another plane. The distinction between complexity and visibility HDAs has not been made before, as far as we know.

A second class of cases that seem to contradict the standard claims about complexity HDAs involves the difference between primary and secondary stress. The troublesome cases are those in which primary stress is quantity-insensitive, whereas secondary stress is quantity-sensitive. We have appealed to a theory of stress that places primary and secondary stress in different planes. Being relatively independent, differences between head feet and dependent feet do not bear on the HDA syndrome.

We have also shown that complexity HDAs can be identified in the

[25] One might ask if the various HDA effects we have discussed simply follow from this acquisition strategy. Though it would be tempting to try to derive HDA effects from acquisition, such a hypothesis does not account for cases where a grammatical process enforces an HDA, or prevents an anti-HDA. Such HDAs appear to involve active principles of grammar. These principles are consistent with the hypothesised acquisition strategy, but do not reduce to it. We thank Bill Idsardi for raising this issue.

phonology of signs, and we have argued that the HDA-syndrome relates directly to the process of language acquisition.

Finally, the present account of HDAs (including apparent counter-examples) is firmly grounded in an explicit theory of phonological representations, involving the fundamental concepts of heads and dependents, and in a principled view of the organisation of grammar, involving different planes representing different types of information.

REFERENCES

Anderson, J. M. & C. J. Ewen (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Battison, R. (1974). Phonological deletion in ASL. *Sign Language Studies* **5**. 1–19.

Beckman, J. N. (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology* **14**. 1–46.

Beckman, J. N. (1998). *Positional faithfulness*. PhD dissertation, University of Massachusetts, Amherst.

Brentari, D. (1990). *Theoretical foundations of American Sign Language phonology*. PhD dissertation, University of Chicago.

Capell, A. & H. E. Hinch (1970). *Maung grammar*. The Hague: Mouton.

Chung, S. (1983). Transderivational relationships in Chamorro phonology. *Lg* **59**. 35–66.

Dresher, B. E. (1981a). Metrical structure and secondary stress in Tiberian Hebrew. *Brown University Working Papers in Linguistics* **4**. 24–37.

Dresher, B. E. (1981b). Accentuation and metrical structure in Tiberian Hebrew. *MIT Working Papers in Linguistics* **3**. 180–208.

Dresher, B. E. (1982). Accentuation and metrical structure in Tiberian Hebrew. *NELS* **12**. 75–85.

Dresher, B. E. (1994). The prosodic basis of the Tiberian Hebrew system of accents. *Lg* **70**. 1–52.

Dresher, B. E. (to appear). Charting the learning path: cues to parameter setting. *LI* **30**. 27–67.

Dresher, B. E. & H. G. van der Hulst (1993). Head–dependent asymmetries in phonology. *Toronto Working Papers in Linguistics* **12:2**. 1–17.

Dresher, B. E. & H. G. van der Hulst (1995). Head–dependent asymmetries in phonology. In van der Hulst & van de Weijer (1995). 401–431.

Dresher, B. E. & A. Lahiri (1991). The Germanic foot: metrical coherence in Old English. *LI* **22**. 251–286.

Dyck, C. (1995). *Constraining the phonology–phonetics interface*. PhD dissertation, University of Toronto.

Emonds, J. (1976). *A transformational approach to English syntax*. New York: Academic Press.

Ghini, M. (1993a). *Phonological phrase formation in Italian*. MA thesis, University of Toronto.

Ghini, M. (1993b). Phi-formation in Italian: a new proposal. *Toronto Working Papers in Linguistics* **12:2**. 41–79.

Grijzenhout, J. (1992). Finnish secondary stress and the foot parsing locality parameter. Ms, OTS, Utrecht.

Hale, K. & E. Selkirk (1987). Government and tonal phrasing in Papago. *Phonology Yearbook* **4**. 151–184.

Halle, M., J. W. Harris & J.-R. Vergnaud (1991). A reexamination of the stress erasure convention and Spanish stress. *LI* **22**. 141–159.

Halle, M. & J.-R. Vergnaud (1987). *An essay on stress*. Cambridge, Mass.: MIT Press.

Hammond, M. (1986). The obligatory-branching parameter in metrical theory. *NLLT* **4**. 185–228.

Harris, J. (1990). Segmental complexity and phonological government. *Phonology* **7**. 255–300.

Hayes, B. (1980). *A metrical theory of stress rules*. PhD dissertation, MIT. Distributed 1981, Indiana University Linguistics Club.

Hayes, B. (1984). The phonology of rhythm in English. *LI* **15**. 33–74.

Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (eds.) *Rhythm and meter*. Orlando, Fl.: Academic Press. 201–260.

Hayes, B. (1995). *Metrical stress theory : principles and case studies*. Chicago: University of Chicago Press.

Hoekstra, T., H. G. van der Hulst & M. Moortgat (1980). Introduction. In T. Hoekstra, H. G. van der Hulst & M. Moortgat (eds.) *Lexical grammar*. Dordrecht: Foris. 1–48.

Hulst, H. G. van der (1984). *Syllable structure and stress in Dutch*. Dordrecht: Foris.

Hulst, H. G. van der (1993). Units in the analysis of signs. *Phonology* **10**. 209–242.

Hulst, H. G. van der (1994). Radical CV phonology: the locational gesture. *UCL Working Papers* **6**. 439–478.

Hulst, H. G. van der (1995). Dependency relations in the phonological representation of signs. In H. Bos & T. Schermer (eds.) *Sign language research 1994*. Hamburg: Signum. 11–38.

Hulst, H. G. van der (1996a). Radical CV phonology: the segment–syllable connection. In J. Durand & B. Laks (eds.) *Current trends in phonology : models and methods*. Salford: ESRI. 333–361.

Hulst, H. G. van der (1996b). On the other hand. *Lingua* **98**. 121–143.

Hulst, H. G. van der (1996c). Separating primary accent and secondary accent. In H. G. van der Hulst, R. Goedemans & E. Visch (eds.) *Stress patterns of the world*. The Hague: Holland Academic Graphics. 1–26.

Hulst, H. G. van der (to appear a). Word accent. In van der Hulst (to appear b).

Hulst, H. G. van der (ed.) (to appear b). *Word prosodic systems in the languages of Europe*. Berlin & New York: Mouton de Gruyter.

Hulst, H. G. van der & G. Rowicka (1997). On some parallels between (un)realised empty nuclei and (un)stressed syllables. In G. Booij & J. van de Weijer (eds.) *Phonology in progress – progress in phonology*. The Hague: Holland Academic Graphics. 125–149.

Hulst, H. G. van der & W. Sandler (1994). Phonological theories meet sign language: two theories of the two hands. *Toronto Working Papers in Linguistics* **13:1**. 43–73.

Hulst, H. G. van der & J. M. van de Weijer (eds.) (1995). *Leiden in last*. The Hague: Holland Academic Graphics.

Hyman, L. M. (1985). *A theory of phonological weight*. Dordrecht: Foris.

Janis, N. (1987). *A grammar of the Biblical accents*. PhD dissertation, Harvard University.

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

Kaye, J., J. Lowenstamm & J.-R. Vergnaud (1985). The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook* **2**. 305–328.

Kaye, J., J. Lowenstamm & J.-R. Vergnaud (1990). Constituent structure and government in phonology. *Phonology* **7**. 193–231.

Liberman, M. & A. Prince (1977). On stress and linguistic rhythm. *LI* **8**. 249–336.

Lightfoot, D. (1989). The child's trigger experience: degree-0 learnability. *Behavioral and Brain Sciences* **12**. 321–375.

McCarthy, J. & A. Prince (1990). Foot and word in prosodic morphology: the Arabic broken plural. *NLLT* **8**. 209–283.

McCarthy, J. & A. Prince (1995). Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers* **18**. 249–384.
Nespor, M. & I. Vogel (1986). *Prosodic phonology*. Dordrecht: Foris.
Newton, B. (1972). *The generative interpretation of dialect: a study of Modern Greek phonology*. Cambridge: Cambridge University Press.
Padden, C. A. & D. M. Perlmutter (1987). American Sign Language and the architecture of phonological theory. *NLLT* **5**. 335–375.
Pater, J. (1994). On the nonuniformity of weight-to-stress and stress preservation effects in English. Ms, McGill University.
Prince, A. S. (1983). Relating to the grid. *LI* **14**. 19–100.
Prince, A. S. & P. Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder.
Rizzi, L. (1989). On the format for parameters [commentary on Lightfoot]. *Behavioral and Brain Sciences* **12**. 354–356.
Roca, I. M. (1986). Secondary stress and metrical rhythm. *Phonology Yearbook* **3**. 341–370.
Ross, J. R. (1973). The Penthouse Principle and the order of constituents. In C. Corum, T. C. Smith-Stark & A. Weiser (eds.) *You take the high node and I'll take the low node: papers from the Comparative Syntax Festival*. Chicago: Chicago Linguistic Society. 397–422.
Sandler, W. (1994). Hand in hand: the roles of the nondominant hand in Sign Language Phonology. *The Linguistic Review* **10**. 337–390.
Sandler, W. (1995). Markedness in handshapes of signs: a componential analysis. In van der Hulst & van de Weijer (1995). 369–399.
Schane, S. A. (1984). The fundamentals of particle phonology. *Phonology Yearbook* **1**. 129–155.
Selkirk, E. O. (1972). *The phrase phonology of English and French*. PhD dissertation, MIT. Distributed 1982, Indiana University Linguistics Club.
Selkirk, E. O. (1986). On derived domains in sentence phonology. *Phonology Yearbook* **3**. 371–405.
Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (eds.) *Studies of child language development*. New York: Holt, Rinehart & Winston. 175–208.
Steriade, D. (1994). *On defining potentially contrastive gestures*. Ms, UCLA.
Stokoe, W. C., Jr. (1960). *Sign language structure: an outline of the visual communication systems of the American deaf*. *Studies in Linguistics Occasional Papers* **8**.
Stowell, T. (1979). Stress systems of the world, unite! *MIT Working Papers in Linguistics* **1**. 51–76.
Trommelen, M. & W. Zonneveld (to appear). Dutch. In van der Hulst (to appear b).
Trubetzkoy, N. S. (1969). *Principles of phonology*. Translated by C. A. M. Baltaxe. Berkeley: University of California Press.
Vergnaud, J. R. (1977). Formal properties of phonological rules. In R. E. Butts & J. Hintikka (eds.) *Basic problems in methodology and linguistics*. Dordrecht: Reidel. 299–318.
Vergnaud, J.-R. & M. Halle (1979). Metrical phonology: a fragment of a draft. Ms, MIT.
Vogel, I. & S. Scalise (1982). Secondary stress in Italian. *Lingua* **58**. 213–242.
Weijer, J. M. van de (1994). *Segmental structure and complex segments*. PhD dissertation, University of Leiden.
Wexler, K. & P. W. Culicover (1980). *Formal principles of language acquisition*. Cambridge, Mass.: MIT Press.
Yip, M. (1989). Contour tones. *Phonology* **6**. 149–174.
Zec, D. (1988). *Sonority constraints on prosodic structure*. PhD dissertation, Stanford University.