On Some Deep Structural Analogies between Syntax and Phonology

Marcel den Dikken & Harry van der Hulst

1 Introduction*

A commonly held view in theoretical linguistics is that the formal organization of phonology is fundamentally different from that of syntax. Claims to that effect in the literature concern either representational aspects or derivational ones (cf. Halle & Bromberger 1989: phonology has extrinsic rule ordering, syntax does not). In the representational domain, it is customary to state that whereas recursion is a fundamental property of syntax, phonological structure is non-recursive:

> "Recursion consists of embedding a constituent in a constituent of the same type, for example a relative clause inside a relative clause (….). This does not exist in phonological structure: a syllable, for instance, cannot be embedded in another syllable." (Pinker and Jackendoff 2005:10)

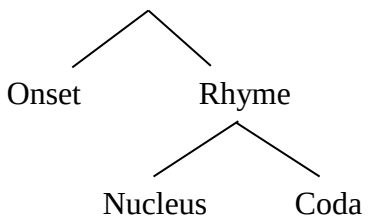
> "syntax has recursive structures, whereas phonology does not." (Neeleman and van de Koot 2006:1524)

> "syllabic structure is devoid of anything resembling recursion." (Bickerton 2000)

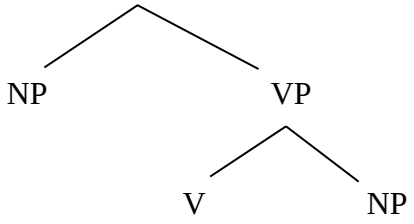
Neeleman & van de Koot (2006:1524), as well as Scheer (2013), even reject the idea that phonological organization appeals to any notion of constituency; see also Carr (2006) for skepticism regarding syntax/phonology parallelism.

Contrary to these views, it has been remarked more than once that there is an 'obvious' parallelism between the structure of syllables (with an onset, rhyme division, and a division between nucleus and coda in the latter) and the structure of a 'simple' sentence (Kuryłowicz 1948; Pike and Pike 1947; Fudge 1987):

(1) a.



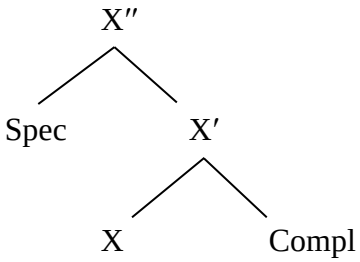
b.



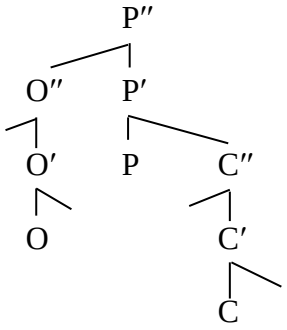
Carstairs-McCarthy (1999) speculates that the structure of syllables may have served as a model for syntax in the course of language evolution, a view that is criticized in Tallerman (2006), who also doubts that the parallelism is real, let alone that syntax copied phonology; Bickerton (2009), in his review of Carstairs-McCarthy (1999), shares this latter view. Despite these objections, various phonologists have pursued the parallelism in (1), and more

specifically a parallelism between (1a) and canonical X-bar structure (see (2), below).[1] Völtz (1999) proposes an X-bar model for syllable structure that explicitly proposes that both the Onset and the Coda can form maximal projections, as in (3) (where 'O' stands for 'Onset', 'P' for 'Peak' (i.e., Nucleus), and 'C' for 'Coda'):

(2)



(3) Universal structure of the syllable



Whatever the merit of these parallels, no mention is made of a potential further parallelism that would involve **recursion**. To avoid miscommunication and controversy, it will be important to be clear on what is meant by 'recursion'. By 'recursion' is understood the containment/embedding of a complex structure inside a larger complex structure of the same type (see e.g. Van der Hulst 2010a). Thus $[_A$ B $[_C$ D $[_E$ F G]]] instantiates recursion: the complex structure labelled 'A' contains another complex structure of the same type, labelled 'C', which in turn embeds another complex structure of the same type, labelled 'E'. Nodes A, C and E are representationally constituted in the same way: we are dealing with the 'Russian doll' structure characteristic of recursion. Note that when it comes to the question of whether nodes A, C and E are of the same type, what matters is their *geometrical properties* – NOT their label.

---

* We would like to thank three reviewers for their helpful suggestions and challenging questions.

[1] Levin (1985) pursues this idea, although in her conception of phonological X-bar structure, the head nucleus can itself be a branching unit (Levin 1985: 105 ff.), which runs counter to standard X-bar theory. We also note that she admits various levels of adjunction (leading to an iteration of the maximal N″ level; see p. 163) to deal with word-final consonant clusters. We will appeal to adjunction in this work as well.

It may be useful to point out at this early stage in the paper that our adoption of standard X-bar theory in what follows does not put us into conflict with current minimalist syntax. **Error! Main Document Only.**X-bar theory is still a staple of generative syntactic theory today, notwithstanding Chomsky's (1994, 1995: . 4) 'bare phrase structure', which aimed to make the bar levels and possibly also the labels redundant: in current minimalist theorizing, both head/phrase distinctions and node labels continue to play a central role (see Chomsky 2013). X-bar theory is such a useful representational vehicle because it regularises recursion in a particularly simple and transparent way, directly codifying the fundamental phrase structure properties of *endocentricity* and *projection*.

It is certainly imaginable that A, C and E have the same label (in which case we are dealing with 'self-embedding' recursion). But even if A, C and E do not have the same label, the structure [$_A$ B [$_C$ D [$_E$ F G]]] is still recursive. In what follows, 'recursion' is understood in its most inclusive sense.

Most writers, while acknowledging that phonotactic structure is constituency-based (and making reference to X-bar(ish) organization of syllables), propose that phonological (often called 'prosodic') constituency is 'strictly layered', which means that no constituent contains a constituent of the same type. This explicitly bars (self-embedding) recursion. With reference to 'higher' phonological/prosodic structure, recursion *has* been recognized, but here it is then said to reflect the recursive structure of syntax, at least to some extent (Ladd 1996 [2008], Wagner 2005, van der Hulst 2010b, Hunyadi 2010).[2] Limiting recursion in phonology to units that have morpho-syntactic structure is tantamount to saying that no recursion will be found *within* morphemes (or simplex words), where whatever structure exists cannot be a mapping from morphosyntactic structure.

However, some phonologists – whose proposals differ in several ways that will not concern us here – have argued that syllable structure can display recursion (Smith 1999, 2003; Garcia-Bellido 2005; van de Weijer and Zhang 2008; van der Hulst 2010b). Following van der Hulst's (2011) cue, the present chapter will support the idea that syllable structure shares non-trivial properties with syntactic structure (parallels that cannot have been inherited from syntactic phrasing), including, crucially, recursion. We will resolve certain problems that arise for van der Hulst's original proposal, which will lead us to introduce structural properties in syllable structure that mirror aspects of more current versions of syntactic structure, specifically proposing a parallel to the so-called 'light *v*' of current 'minimalist' syntactic inquiry.

Our principal conclusion is that there is only one syntactic (or 'computational') system which underlies both phonological structure and morpho-syntactic structure (as well as operations). Whatever differences are found between the two systems are primarily due to the fact that both modules differ in their basic alphabet. Thus, we support what John Anderson calls 'The Structural Analogy Assumption' (SAA; Anderson 1987):

(4) The Structural Analogy Assumption

> The same structural properties are to be associated with different levels of representation except for differences which can be attributed to the different character of the alphabet involved (as in the case of planes) or to the relationship between the two levels (as may be the case with any pair of levels), including their domains.

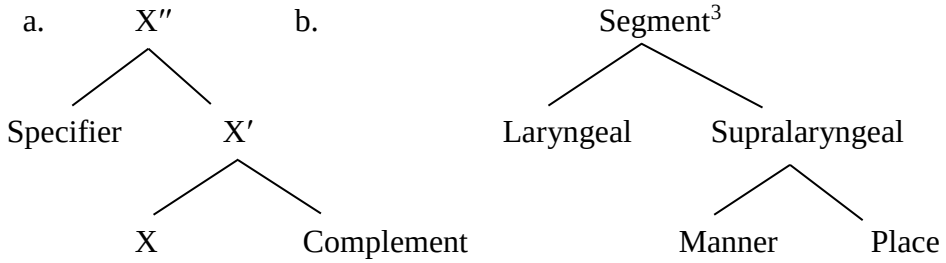
Here 'planes' refers to syntax and phonology. Structural analogy holding between levels within planes will not be our concern here. Anderson pursues the SAA within a dependency framework. van der Hulst (2005, in prep.) develops Anderson's dependency approach in his Radical CV Phonology model. While Anderson works within a dependency model (which, crucially, does not recognize constituency), we examine parallels between syntax and phonology from a headed constituency perspective. We will not dwell on this issue here. Our main thesis ('there is only one syntactic system') can be worked out in different ways depending on the precise

[2] In section 6 we briefly discuss the question as to what limits phonological recursivity in morpho-syntactically structured expressions.

syntactic and phonological structures that we compare. In both domains, there has always been, and will continue to be, development, which, at times, may suggest that there are no analogies at all, or that resemblances are trivial or coincidental. As a consequence, the recovery of pervasive analogies may require presenting structure in one domain or the other in perhaps novel ways, which may lead to new perspectives on the representation in either domain. In this chapter, we take a particular proposal for syntactic structure as our point of departure, showing that parallel structures may shed new light on phonological phenomena.

Though in this chapter our focus will be on syllable structure, we will also address segmental structure, including the potential interweaving of both levels. In this context, we will discuss the applicability of X-bar structure within phonological segments, as in (4b) (van der Hulst 2005):

(5) a. b.



The chapter is organized as follows. In section 2 we will begin with outlining a proposal made in van der Hulst (2010b). Section 3 then develops this idea in a more in-depth, leading to a proposal to import 'light *v*' structures into the representation of syllables and feet. In Section 4, we show how the model accounts for different types of 'foot structure' (trochaic, iambic, coordinate). Section 5 discusses segment-*internal* X-bar structure (cf. 5). Here we also address the issue of 'segmental integrity', i.e. whether segmental structure and syllable structure are strictly separated or rather, as we will argue, integrated. In section 6 we offer an explanation for the fact that recursion in phonology is less pervasive than in syntax. Section 7 offers our main conclusions.

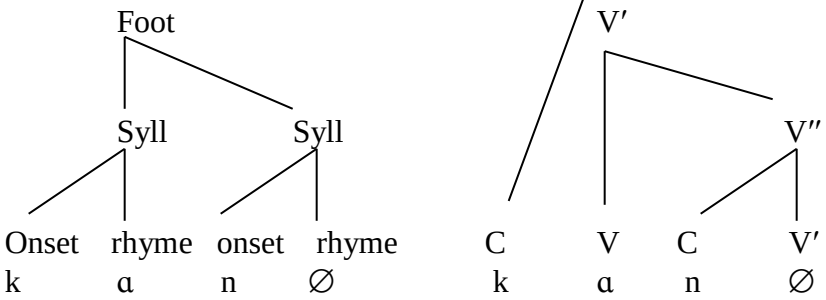
2 van der Hulst (2010)

The central point of van der Hulst (2010) lies in a particular construal of the idea that so-called 'Codas' can be entire syllables. Adopting his 'C/V notation',[4] van der Hulst (2010b) proposes the structure in (6b) rather than the more traditional (6a) for a 'monosyllabic' word like Dutch *kan* 'can'. In approaches such as Government Phonology, especially those versions that adhere to a strict CV principle, such a monosyllabic word would be a sequence of two 'syllables' (or Onset/Rhyme 'packages'), which could then be taken to form a structure or lateral relation comparable to a 'trochaic foot':

[3] This structure follows the original proposal in Clements (1985). van der Hulst (2005) argues that the later idea to abandon a manner node (attaching manner features directly to the root node) should not be followed.

[4] Note that we are not claiming in this paper that all of language is built up from Cs and Vs. This is a specific proposal for phonology. It may be that phonology and syntax could ultimately be tackled with the help of the same two primitives in both domains (and that the labels for these primitives should be different from 'C' and/or 'V'), but this is not under discussion here. This chapter is about representation, not substance.

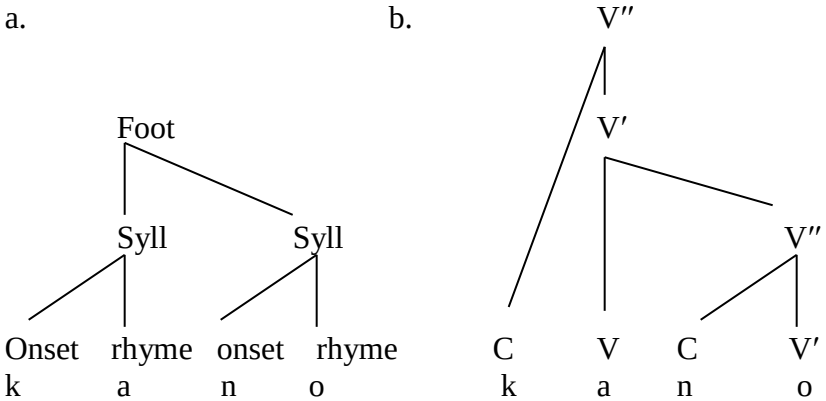
(6) a. b.



In the notation in (5b) the labels 'C' and 'V' are analogous to the labels 'N' and 'V' in syntax; they are phonological categories to which segmental units can be associated. Taking the V unit to be the head of a syllable (which therefore, as a whole, belongs to the category V as well), Codas are complements, which are thus expected to be maximal projections (cf. (2)). The crucial point in (6b) is that the complement of the V-head is a maximal V-projection (in short, a complete syllable).

The next step in van der Hulst's proposal is to adopt the same kind of structure for more obviously disyllabic strings such as Dutch *káno* 'canoe', as in (7). This establishes a perfect isomorphism between a 'closed syllable' and a branching foot, which, from a metrical point of view, behave as units of stress in languages, such as Dutch, in which 'closed syllables' are heavy for stress. This equivalence is widely acknowledged, yet does not find a formal basis in any other model, although so-called moraic models capture the equivalence by referring to the fact that a closed syllables contain two morae, on a par with a sequence of two light syllables.[5]

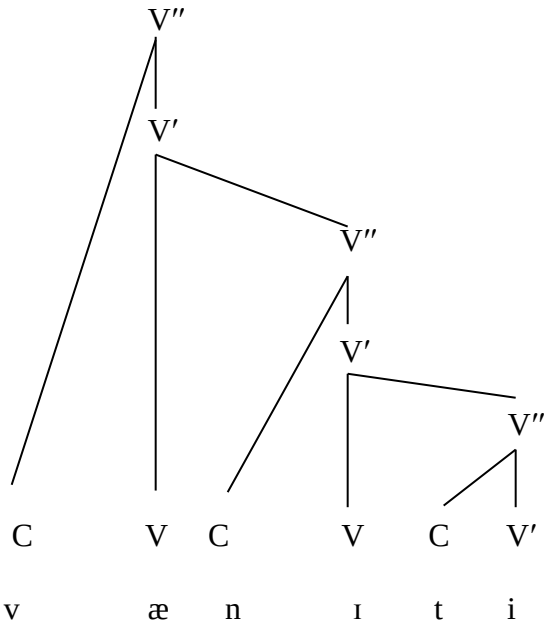
(7) a. b.



The embedding of syllables inside syllables does not have to stop here. A full structure of a so-called ternary foot, sometimes referred to as a 'superfoot' (as in English *vanity*), displays degree-2 embedding.

[5] A different proposal for the structural equivalence between CVC 'heavy syllables' and CVCV (feet) can be found in Ulfsbjorninn (2015) within the 'strict CV' Government Phonology model.

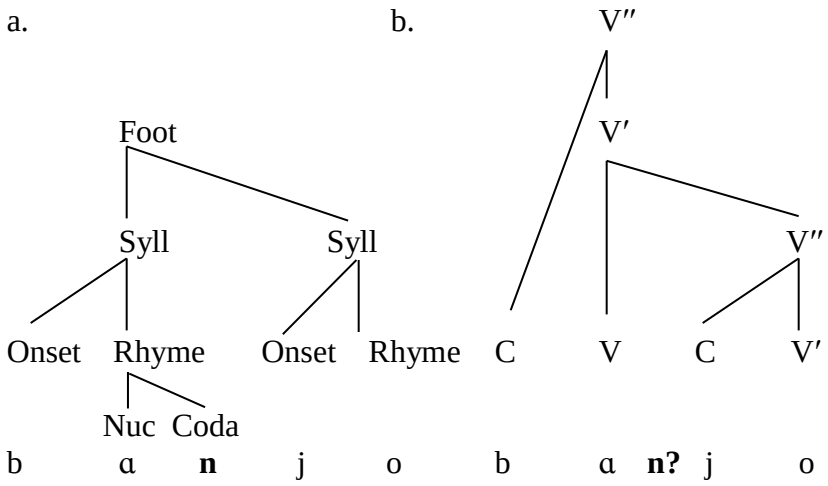
(8)



This structure is a perfectly legitimate object also in, for example, English *Winnepesaukee*, *hippopotamus*. An interesting consequence of this proposal is that it is now immediately clear why in poetic rhyming the initial Onset can be ignored, but not the second (or indeed the third, in forms like *sanity ~ vanity*). The initial consonant of such structures is external to the whole sequence that forms the rhyming unit. The structure in (8) formally captures the rhyming unit as well as the special position of the initial Onset (which can or must be different), as opposed to the other more deeply embedded Onsets (which must be identical).

The preceding proposal faces one problem: a matrix syllable can itself be a 'closed syllable' (as in *banjo*; in (9) we represent the Dutch pronunciation), which would seem to leave no room for the closing /n/ consonant, given that the 'Coda' position in (9b) is taken by the syllable /jo/:

(9) a. b.



Three apparent 'solutions' present themselves, which we will briefly discuss (and dismiss) in the ensuing paragraphs.

Firstly, one might consider adjoining to /n/ to the syllable head, thus forming a branching 'Nucleus'. But this would destroy the X-bar analogy since the head of an X-bar projection must be 'atomic'; it cannot itself contain a full X-bar-theoretic internal structure (there are no phrases within heads).[6]

A second possibility would be to embrace a proposal by Botma, Ewen, and van der Torre (2008), where an analysis is given for a range of facts clustering around postvocalic liquid+stop sequences in English. One of the more striking properties of such sequences is that when they occur after a long tense vowel or diphthong, the stop must be coronal (see (10a)). For postvocalic liquid+stop sequences that occur after a short vowel, no such coronality restriction applies, as (10b) shows.

(10) a.

wield	/wi:ld/	colt	/kəʊlt/
*wielk	*/wi:lk/	*colp	*/kəʊlp/

b.

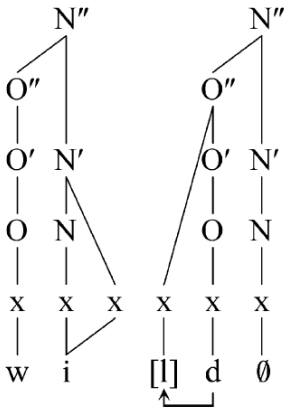
silt	/sɪlt/	Celt	/kɛlt/
silk	/sɪlk/	kelp	/kɛlp/

The well-known generalization that lies behind these data is that tense vowels are equivalent to lax vowel + one consonant. As such both 'exhaust' the bipositional rhyme. Word-finally, bipositional rhymes can be followed by one 'extra' consonant (as in *team* or *film*, where the extra consonant is /m/ in both cases) and 'extra' coronal consonants (traditionally referred to as the 'appendix'; see Fudge 1987), raising the question how these 'extras' are structurally represented.

Botma *et al.* argue that in *wield* the 'extra' liquid and the 'appendix' /d/ can form an onset to a (silent-headed) second syllable, with the liquid being an 'Onset Specifier'. The two consonants are said to end up in a *Specifier-Head agreement* relation, which is taken to account for the fact that the stop (the head of the Onset) and the liquid (the Onset Specifier) will share their place specification:

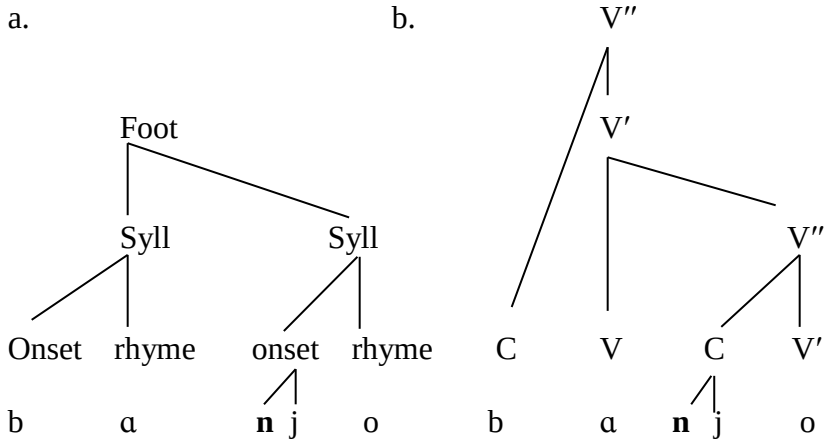
(11) (= Botma *et al.*'s (26))

[6] We ignore at this juncture the segmental-internal complexity of the nuclear vowel itself; we return to this issue in section 5.



Applied to the case at hand, one might consider extending this idea to the structure of *banjo* as follows:[7]

(12) a. b.



A problem with this idea is that sequences such as /ld/ and /nj/ systematically fail to serve as Onsets of word-initial syllables in English and Dutch (i.e., there are no words beginning with *ld* or *nj*). While differences between word-initial and medial onset do exist, it is then usually the case that word-initial onset display more option, not less; e.g. in Dutch /kn/ is possible word-initial, but not word-medially; see Trommelen (1983) and van der Hulst (1984). Also, the appeal made by Botma *et al.* to Specifier-Head agreement to force the stop to be coronal seems to us to be a misapplication of a syntactic notion to phonological analysis. To the extent that relations of Specifier-Head agreement arise in syntax, they are found only in *functional* structures (the head I usually agrees with SpecIP, C sometimes agrees with SpecCP), not in lexical ones ('object agreement' is by no means rare cross-linguistically, but arguably implicates a functional head outside the lexical core); and phonology arguably lacks anything corresponding to functional structure in syntax (see the opening paragraph of section 3.2, below).[8]

[7] They do not consider, nor will we here, postulating an 'empty nucleus' between the /n/ and /j/, arriving at a 'trisyllabic' structure, as would, or could be adopted in a (strict CV) government approach; but see fn.8.

[8] A third issue that arises in connection with Botma *et al.*'s (2008) proposal is that these authors explicitly do not exclude a representation for *wield* in which the liquid (now realized as a 'dark' [ɬ]) is mapped by itself into the Coda position of the second syllable in a trisyllabic sequence, with the stop as the Onset of the third syllable;

A third approach would be to give embedded syllables a 'complementizer' position, which could contain the 'Coda' consonant. But this would again entail an introduction into phonological structure of functional layers, which seems to us unwarranted. We return to this lack of analogy between syntax and phonology in the next section, where we propose to 'enlighten' phonology with the introduction of a parallel to the syntactic notion of 'light *v*'.

3 Enlightened phonology: The benefits of 'light *v*' in phonology

3.1 'Light v' in syntax: A brief historical perspective

In syntax, the external argument of the verb is different in a number of fundamental ways from the verb's internal argument(s). Thus, the internal argument(s) can influence the aspectual (*Aktionsart*) properties of the sentence but the external argument never does; and the verb can form an idiomatic expression together with its internal argument(s) but not with its external argument. Such pervasive asymmetries led Kratzer (1996) to hypothesize that the external argument is *radically* external to the 'minimal VP': it is introduced in the specifier of an extension of the core verbal phrase.[9] Kratzer called this extension 'VoiceP', based on the insight that the external argument is syntactically projected only in certain voices (the active, perhaps the passive, but certainly not the middle voice or *vox media*). Chomsky (1995) bought into the idea that the external argument is severed from the core verbal phrase, and called the extension of VP in whose specifier the external argument is introduced '*v*P', where '*v*' is a 'light verb' merged immediately with VP. This *v* is a lexical head in the sense that it plays a key role in the syntactic deployment of the argument structure of the verb. In this respect, it is fundamentally different from purely functional categories such as I(nfl) or C(omp).

Severing the external argument from V gives the VP more space to accommodate internal arguments, along the lines of Larson's (1988) original proposal for the syntax of ditransitive constructions: with the specifier position of V no longer needed for the introduction of the external argument, it can be used for one of the internal arguments of the verb. Larson (1988) and Hale & Keyser (1993) converge on the conclusion that the SpecVP position, when filled by an argument at D-structure, is reserved for the Theme argument (the argument of which a (change of) state or position is predicated, as in *John broke <u>the vase</u>* and *<u>The vase</u> broke*: in both sentences, *the vase* is introduced in SpecVP; in the second example, it is raised from there to SpecIP). In keeping with the Uniformity of Theta Assignment Hypothesis (UTAH; Baker 1988), which says that identical thematic relations between items are represented by identical structural relations between those items at the level of D-structure,

---

see their (27). Apart from the fact that this creates structural ambiguity, it would seem that there is now nothing about the structure in which the liquid and the stop are mapped into different syllables that could be held responsible for the shared coronality of the liquid and the following stop: the two are not in a Spec-Head agreement relation here.

[9] We note on the side that the external argument is often externalized further, to SpecIP. But relying on such externalization cannot suffice to ensure across the board that the external argument is different from the internal argument(s) in ways that involve argument and event structure: even when the external argument is not raised to SpecIP (as e.g. in transitive expletive constructions such as Dutch *Er at iemand een appel* '(lit.) there ate someone an apple, someone was eating an apple'), it still behaves differently from the internal argument(s) in these ways.

there is a tight connection between base-generation of an argument in SpecVP and the Theme role. The complement-of-V position is used for non-Theme material: an argument projected in this position can be a Patient (as in *John hit Bill*) or a propositional argument (as in *Bob saw* [*that John hit Bill*]); non-arguments (including secondary predicates, such as the *to*-PP in prepositional dative constructions, and on Larson's assumptions even certain adverbial modifiers) can also be merged in the complement-of-V position. Apparently V is rather flexible regarding the relations between itself and its complement.[10] The UTAH has always been most successful with respect to predicate–argument relations involving specifier positions: SpecVP is tied one-to-one to the Theme role, Spec*v*P is usually (and, depending on one's approach to sentences such as *John fears snakes*, with an Experiencer subject, perhaps exclusively) tied to the Agent role. The link between the complement-of relation and thematic roles is much more obscure. Until more is known about the thematic properties of the complement-of relation, it will be sensible to confine the scope of the UTAH to specifiers:

(13) *Uniformity of Theta Assignment Hypothesis (UTAH)* (specifier-only version)
Specifier positions in the lexical core[11] are associated with unique thematic content in underlying representations.

Interestingly, as soon as the complement-of-V position is taken, an additional argument of the verb that is not its external argument must be projected in SpecVP, and will, in keeping with (13), necessarily be construed as a Theme. This explains straightforwardly that *John hit Bill* and *John hit Bill unconscious* are minimally different with respect to the thematic role of *Bill*: in the former sentence, where *Bill* is in the complement-of-V position, *Bill* is the Patient; in the latter, with the secondary predicate *unconscious* now occupying the complement-of-V position, *Bill* must be mapped into the SpecVP position, and is interpreted as the Theme.

To summarize, the projection of the 'light verb' *v* in syntax is an extension of the lexical VP and harbors the external argument (Agent) of the verb, freeing up the SpecVP position for the projection of the Theme, and the complement-of-V position for the introduction of non-Theme dependents of the verb. The range of possible verb phrases with an external argument that the 'light verb' hypothesis gives rise to can be summarized as follows:[12]

---

[10] We see this flexibility also in our discussion of the structure of the phonological *v*-VP: the complement-of-V position in phonology can be filled by a variety of different consonant types, and by 'propositional arguments' (i.e., dependent syllables in trochaic feet).

[11] By 'the lexical core' in syntactic structures, we mean minimally the root-VP, perhaps plus its 'light' extension *v*P but excluding functional projections higher up the tree. It should be noted that the most successful applications of UTAH have always been focused on the relation between SpecVP and the Theme role, which does indeed seem to be very strict: any argument externally merged in SpecVP is a Theme. For the external argument, things are less clear cut: much depends on how microscopic one's view of the structure of the lexical core outside VP is. Observationally, external arguments of verbal constructs can be Agents (as in *John hit Bill*) or Cause(r)s (as in *The earthquake destroyed the village* or *John accidentally broke the vase*) or Experiencers (as in *John fears snakes*). Different flavors of *v* can be introduced to differentiate between Agents and Cause(r)s, and Experiencers could possibly be introduced as internal arguments and externalized via raising (cf. *Snakes frighten John*). Our focus here, as in the discussion of syllable structure, will be on the restrictions imposed on SpecVP.

[12] In all of these structures, 'DP' stands for 'Determiner Phrase' (the 'noun phrase' including any and all of its functional attributes), and the subscript on DP references the thematic role borne by the argument in question. In (14f), 'SC' stands for 'small clause'. In addition to these verb phrase types, there may also be the possibility of not projecting *v* and, as a consequence, not having an external argument. This may be what characterizes the syntax of unaccusative/ergative constructions. We will not need to concern ourselves with these here.

(14)	a.	$[_{vP}\ DP_{Agent}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ V]]]]$	*John laughed*
	b.	$[_{vP}\ DP_{Agent}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ V\ DP_{Patient}]]]]$	*John hit Bill*
	c.	$[_{vP}\ DP_{Agent}\ [_{v'}\ v\ [_{VP}\ DP_{Theme}\ [_{V'}\ V]]]]$	*John killed Bill*
	d.	$[_{vP}\ DP_{Agent}\ [_{v'}\ v\ [_{VP}\ DP_{Theme}\ [_{V'}\ V\ Pred]]]]$	*John hit Bill unconscious* *John hit Bill into the hospital* *John gave a book to Bill*
	e.	$[_{vP}\ DP_{Agent}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ V\ CP]]]]$	*Bob saw that John hit Bill*
	f.	$[_{vP}\ DP_{Agent}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ V\ [_{SC}\ DP\ Pred]]]]]$	*Bob saw John hit Bill*

The *v*-VP structure forms an integral part of syntactic structures. The 'light verb' extension of the lexical VP is a key ingredient in our understanding of thematic relations as well as event structure and aspect. To this *v*-VP structure, adverbial material can be adjoined, and outside it functional projections can be introduced, such as IP and CP, whose role it is to regulate properties of syntactic constructs that are not the purview of the argument-structural core, such as inflection, negation, mood and modality, question formation, and variation in linearization.

3.2 'Light v' in phonology: Preliminary remarks

Throughout this work, we explore the possibility that phonology projects X-bar structures entirely analogous to those recognized in syntax. To be sure, phonological and syntactic structures are not fully on a par: arguably, phonology entirely lacks the kinds of *functional* projections to which we alluded at the end of the previous paragraph. Phonological structures employ the basic 'argument-structural' layers and their complement and specifier positions, and they may also make fruitful use of the adjunction operation to bring in additional material that cannot be accommodated in the complement and specifier positions in the *v*-VP structure.[13] But there is no obvious role to play in phonology for functional projections such as IP and CP, whose specifier positions are usually not filled by base-generation but get occupied as a result of movement operations that externalize material from the argument-structural core of the structure. Phonology provides no plausible cases of such externalization: melodic material is always associated with positions internal to the core. The kinds of long-distance dependencies seen in raising and operator movement constructions in syntax, for which an appeal to specifier positions of functional projections is called for, are entirely absent from phonology, as are candidates for exponence of the heads of such functional projections (i.e., phonology has no plausible counterparts to such staples of syntactic constructs as determiners or complementizers). Like complex morphological constructs, phonological structures arguably lack functional structure altogether; functional structure is the province of the kinds of dependencies that syntax specializes in. In part, functional categories are licensers of properties which cannot be satisfied in the position of External Merge (such as case and agreement, or [+wh]). For the remaining part, functional categories are present in order for syntax to get a handle on variation in linear order involving information-structural properties (topicalisation, focalisation, extra-

[13] In our discussion of foot structures in section 4 and segmental structure in section 5, we will make an appeal to adjunction in phonological X-bar structures. Our focus until then will be on specification and complementation.

position, etc.). Neither of these considerations comes into the picture in phonology. This is why functional structure has no place in phonology.

But though phonology arguably does not deal in functional categories projecting outside the core, one of our major claims in this chapter is that it does recognize the same kind of complex representation of the core that syntax has been argued to feature: on top of the projection of V (which in phonology represents the *vowel*, not the *verb*) we will have occasion to postulate a projection of a 'light *v*'. In syllabic structure, it is the specifier position of *v*P that harbors the Onset, which is the analogue of the external argument in syntax. Inside VP, the structure of the syllable accommodates a variety of different material, often but not invariably associated with the traditional Coda constituent. We will discover that there are interesting regularities regarding the association of melodic material with the SpecVP position in the structure of the syllable – regularities that are reminiscent of those discovered for syntax under the rubric of UTAH. Thus, we announce the birth of a phonological cousin for UTAH, which we will name UMAH:

(15) *Uniformity of Melody Assignment Hypothesis (UMAH)*
Specifier positions[14] in the syllabic core[15] are associated with unique melodic content in underlying representations.

One UMAH subgeneralization that will emerge from the discussion to follow is that whenever the syllabic SpecVP position is underlyingly associated with melodic content, this content must be *sonorant*: non-sonorant material cannot be mapped into SpecVP in underlying representations. This corresponds, as we will see, to the observed tendency for Coda consonants to be (restricted to) sonorant consonants, with SpecVP corresponding to one of the structural positions that can be mapped into the traditional Coda.

Obviously, and superficially, coda consonants in many languages can be non-sonorant, but in such cases, as we argue below, these obstruents are located in the complement-of-V position, which, unlike SpecVP,is not limited to sonorants.

Another interesting property of SpecVP in phonology which we will discover is that the presence of this position is required when we are dealing with a lax vowel, which is spelled out in *v* (i.e., is a 'lax vowel'): what this suggests is that lax vowels are like 'affecting verbs' (verbs that always take a Theme argument, projected in SpecVP).

---

[14] The formulation here refers to positions, but in the present work its application is only rigorous for SpecVP, which is tied to sonorant. Since this position is flanked by little *v* and V (which are both vowel positions and thus sonorant), one might argue that this is why SpecVP, sandwiched between two sonorant elements, must also be sonorant. We note, however, that the SpecVP position can also become associated with non-sonorant melodic content, via Internal Merge: see the discussion in section 3.4. Just as in syntax (see (13)), association of content with core specifier positions is restricted only in underlying representations (i.e., for cases of External Merge).

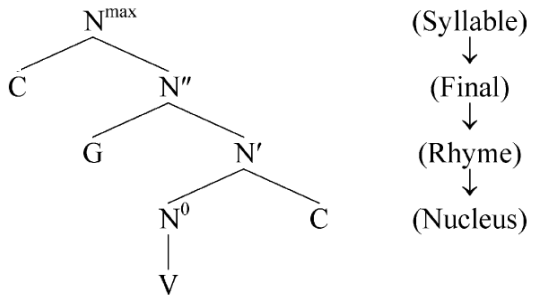
[15] As in the case of 'the lexical core' in our syntactic discussion, by 'the syllabic core' in phonological structures we mean minimally the root-VP, perhaps plus its 'light' extension *v*P. And once again our focus will be on SpecVP, which seems privileged to accept only sonorant material under External Merge. The complement-of-V position is clearly tolerant of a wide range of different constituents (consonantal as well as 'propositional': entire syllables can be embedded in the complement-of-V position, as we will see). The specifier position of *v*P is reserved for consonantal material, but its melodic specification seems much more variable than the melodic specification of SpecVP. Here again there is a parallel with syntax (recall from fn. 2 the range of theta-roles that external arguments can have).

With these remarks as background, let us now develop the *v*-VP structure of the syllable and the roles played by the 'light *v*', and highlight some of the salient benefits of this structure.

3.3 Chinese prenuclear glides in the v-*VP structure of the syllable*

Van de Weijer & Zhang (2008) tackle Chinese prenuclear glides with the help of a syntax-inspired 'X-bar structure' with multiple specifiers, such that the glide is in the inner specifier position and the onset in the outer specifier:

(16) (= Van de Weijer & Zhang's (18))



The main point of this structure is that it allows the authors to express the fact (which they demonstrate in detail) that Chinese prenuclear glides belong neither to the Onset nor to the traditional Rhyme.

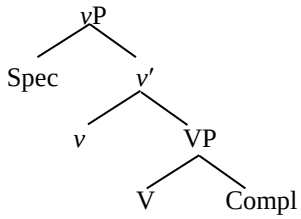
Note, however, that to say that (16) is an X-bar-theoretic approach to Chinese prenuclear glides presupposes a major modification of traditional X-bar theory: multiple specifiers are not available in X-bar theory proper; in fact, it is only when one abolishes traditional X-bar theory (as in Chomsky's 1994, 1995: ch. 4 'bare phrase structure') that multiple specifiers become available.[16] Also, accommodating Chinese prenuclear glides with the aid of a structure of the type in (16) does little to alleviate the Chinese-specific nature of the analysis: multiple specifier structures of the type in (16) do not seem to have any demonstrated or apparent use outside the realm of prenuclear glides in Chinese. Let us therefore explore a different approach, one which eschews multiple specifiers, and exploits X-bar structures familiar from current syntactic analysis.

In developing our analysis of Chinese prenuclear glides, we take optimal advantage of the hypothesis that the V-projection in phonology (for the vowel) can be associated with a structural extension projected by a 'light *v*', just as the V-projection in syntax (for the verb) can have a 'light *v*' on top of it:

---

[16] Note that the simplified X-bar-theoretic structures of Kayne (1994), which do away with the X′/XP distinction, do not allow for multiple specifiers any more than traditional X-X′-XP structures do. It is really only the complete abolition of traditional X-bar labels that makes multiple specifier structures legitimate. The usefulness of such structures in syntactic analysis has always remained a controversial matter. Thus, for multiple nominative constructions in Japanese (which served as the typical illustration of a TP with multiple specifiers) analyses are available which do not require any particular functional head to accommodate more than a single specifier. Chomsky's (1994, 1995: Ch. 4) introduction of multiple specifier structures was born out of the desire to allow *v* to both introduce the external argument in a specifier position and check accusative Case against the object in a specifier position. In more recent approaches, Case is checked under Agree, and 'object shift' no longer targets Spec*v*P.

(17)



The specifier position of the *v*-projection is the position for the 'traditional' Onset; the complement position of V is the standard position for the traditional Coda, although, as we will see in the ensuing subsections, the option of a SpecVP position (which is not shown in (17)) can accommodate melodic material that falls under the traditional notion of Coda; this would be the site for sonorant consonants mentioned above. When *v* and V are spelled out together, and realized ('spelled out') at *v*,[17] the *v*′ represents the traditional Rhyme; when spell-out of *v*-V is at V, it is VP that corresponds to the Rhyme. On the basis of the structure in (17), then, the Rhyme is defined as the minimal structural constituent containing the spell-out position of the syllable Nucleus and its complement (if any).

We propose that Chinese syllables containing a prenuclear glide are characterized by the fact that the glide spells out the *v*-position in the structure in (16), and the Nucleus is spelled out at V – in other words, *v* and V get discrete lexicalizations; we are dealing with a 'serial vowel construction', parallel to 'serial verb constructions' in syntax (for which at least a subset is plausibly analyzed as *v*-V sequences in which *v* and V are spelled out separately; see e.g. den Dikken & Sybesma 1998). When *v* is spelled out as a glide and the syllable nucleus is spelled out at V, the Rhyme corresponds to VP (because V is spelled out); and the Onset of course remains the constituent in Spec*v*P. The prenuclear glide sits right in between the Onset and the Rhyme, and does not strictly belong to either – though, to be sure, it *is* the head of a structural extension (the *v*P 'shell') of the nucleus.[18]

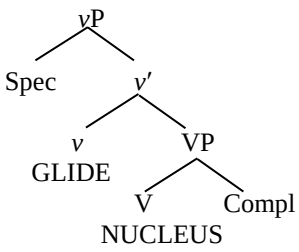
---

[17] Let us clarify what we mean by 'spell(ing) out'. In all cases in which *v* does *not* have melodic content different from that of V (thus unlike what we saw in the Chinese case), the *v* and V positions enter into a chain (cf. 'head movement' in syntax). This chain, which has its melodic content contributed by V, needs to be spelled out in one of the two positions tied together by the chain. In the default case, spell-out of melodic content is at V; but as we will see in our discussion of the tense/lax distinction in section 3.4, spell-out at *v* is what characterizes lax vowels.

[18] Note that this is not tantamount to claiming that the prenuclear glide, by itself, is the head of the syllable. The head of the syllable is the *v*-V complex. In Chinese words featuring a prenuclear glide, the two parts of this complex each have their own surface exponent: the glide spells out *v*, and the vowel is the exponent of V. (For the English diphthong /ɪə/ in words like *weird*, it also seems plausible to say that *v* and V have discrete exponents, /ɪ/ and /ə/, resp.; again, the head of the syllable is the *v*-V complex as a whole. See the discussion of (34b) in section 3.5.)

As a logical alternative, the Mandarin prenuclear glide could in principle be accommodated in SpecVP, with the Rhyme then confined to the V′ node. But because we are dealing, in the Mandarin cases, with a glide that is transparently vocalic in origin, it seems to us more attractive to place this glide in a vocalic position: *v* in the structure in the main text. Moreover, on our approach, the Rhyme can be defined as a maximal projection: VP. In virtue of the fact that both *v* and V have melodic content, a Mandarin syllable with a prenuclear glide is – on the representation in the main text – a kind of sesquisyllable (i.e., a syllable and a half).

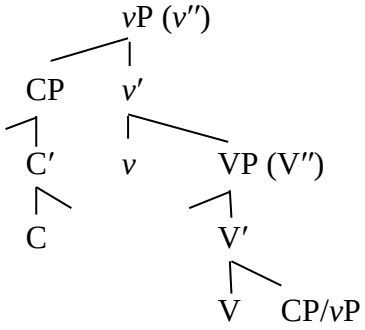
(18)



From this perspective, the parameter that distinguishes Chinese from, say, English when it comes to prenuclear glides is that whereas in English they are an integral part of the Onset (which can have its own X-bar structure, thus allowing for multiple consonants; see (19)), in Chinese these glides are lexicalizations of *v* (in between the Onset and the spell-out site of the Nucleus, i.e., V) – a possibility afforded by the license to spell out *v* and V by discrete elements. This kind of parametric difference between languages resembles the parametric difference in syntax between serializing and non-serializing languages: languages that have 'serial verb constructions' allow *v* and V to be spelled out by different elements whereas languages that do not will lack such constructions. (In point of fact, it turns out that Chinese not only has 'serial *vowel* constructions' (i.e., prenuclear glides) but also 'serial *verb* constructions' – but this is probably an accident rather than something 'deep': we see no particular reason to expect that the 'serialization parameter' will be set the same way for vowels and verbs within individual languages.)

The proposal in (18) embodies what we present as the universal structure of the syllable, encapsulated in (19), which replaces Völtz's (1999) structure in (3):

(19) Universal structure of the syllable



As shown, the complement of V can either be a 'consonant phrase' or, indeed, a full syllable (which is how we incorporate van der Hulst's proposal that syllables can contain syllables, which we will return to below).

3.4 The tense/lax distinction in the v-*VP structure of the syllable*

The benefits of the *v*-VP structure of the syllable also come to the fore in the representation of the difference between long vowels and diphthongs, on the one hand, and short vowels, on the other, and in distinguishing tense and lax vowels. In this section, we will talk about the latter.

Our central hypothesis regarding the difference between tense and lax vowels is (20):[19]

(20) a. lax vowels are spelled out at *v*
b. tense vowels are spelled out at V

We will use this hypothesis in an analysis of the Dutch vowel system and syllable structure.

Dutch systematically distinguishes in its phonology between two types of vowels, often differentiated by the labels *tense* and *lax*. Of these, the former are often phonetically 'long', but extra duration is not the unifying property of tense vowels – not all tense vowels are phonetically long (in particular high tense vowels are quite short), and open syllables with a tense vowel count as light (not heavy) in terms of stress (see van der Hulst 1984), which supports the claim that they are not phonologically long.[20] The lax vowels are marked; in concert with this, Dutch has fewer lax vowels than tense vowels. Lax vowels must be followed by a consonant; tense vowels do not have to be, but when they are (largely only in word-final position), they deliver so-called 'superheavy' syllables. Word-internally, tense vowels tend not to be followed by a tautosyllabic consonant (while lax vowels must be followed by a consonant, arguably tautosyllabic; see the discussion of *kop* 'cup' and *kom* 'bowl' later in this section). Word-finally, where extra consonantal material is possible (see below), tense vowels can be followed by one consonant less than lax vowels. These are the main explananda. In the following paragraphs, we will show that the *v*-V system provides insightful explanations for them, and establishes interesting parallels with the structure of the syntactic verb phrase.

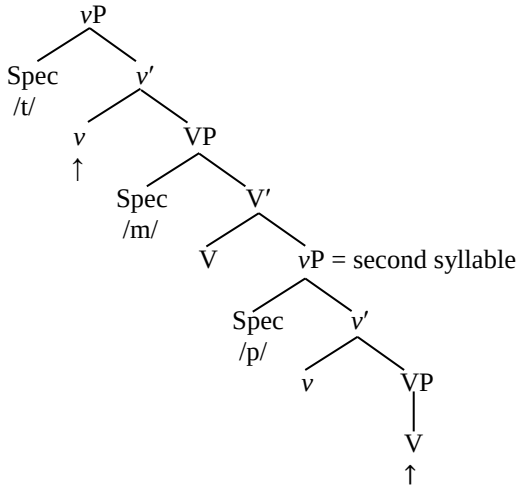
For the contrast between /tɛmpo/ *tempo* 'id.' and */tempo/, the central hypothesis in (20) gives us an immediate account, in conjunction with our previous hypothesis that sonorant consonants can be mapped into SpecVP: see (21) (where, as before, the arrows point to the spell-out position for the Nucleus). With the tense vowel spelled out at V, the b-structure in (21) cannot yield */tempo/.[21] What we see here is that the complement of V can not only be CP (as, for instance, in (22), above), but also *v*P. This testifies to the 'flexibility' of the complement-of-V position, and captures the van der Hulst proposal for syllable embedding, but it avoids the problem discussed in section 2 because the presence of an embedded syllable does not come in the place of a closing consonant for the first syllable, which occurs in Spec of VP (while it is still true that the embedded syllable vP and CP are in complementary distribution as options for the complement of V.

[19] This hypothesis was ultimately inspired by Polgárdi's (2008) approach to the tense/lax distinction in Dutch, although the two outlooks differ fundamentally. Polgárdi's idea that Dutch lax vowels *must* properly govern a silent Nucleus strikes us as an anomaly: proper government is always a privilege, never an obligation.

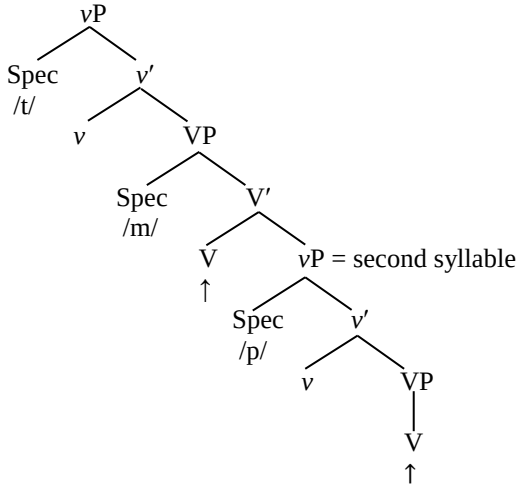
[20] If they were phonologically long, given that stress is Dutch is weight-sensitive, these vowels would be expected to attract stress, but they do not.

[21] Spelling the b-structure in (21) out as is would deliver the ungrammatical */tmepo/, which is bad because /tm/ does not occur tautosyllabically in prevocalic position. Dutch proper onsets (as occurring word-internally) cannot have a nasal in second position. In the text discussion above (27), we argue that a tense vowel, spelled out at V, cannot have SpecVP filled with independent melodic content – and we link this to a kind of 'doubly-filled Comp effect'.

(24) a. /tɛmpo/ *tempo* ‘id.’

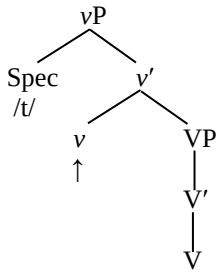


b. */tempo/

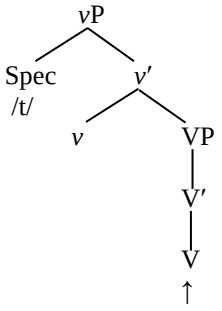


The lax vowel /ɛ/ is legitimate as the head of the stressed initial syllable in (21a) thanks to the fact that there is a Coda consonant present in SpecVP: the nasal /m/, a sonorant that is eligible for insertion in SpecVP. We will explain below what happens when a stressed lax vowel is followed by a non-sonorant consonant. But let us first discuss what goes awry when a lax vowel heads a stressed *open* syllable, as in (22a), to be contrasted with (22b).

(22) a. */tɛ/



b. /te/ *thee* ‘tea’



Without any segmental material in SpecVP, the distinction between *v* and V as the spell-out site for the Nucleus cannot be made: shifting the arrow from V to *v* would be a vacuous operation. Universally, the default spell-out position for the Nucleus is V. Because shifting the spell-out site of the vowel over to *v* is vacuous in the absence of an occupant of SpecVP, it follows that without any segmental material in SpecVP, (22) can only be realized as /te/, with the vowel spelled out at V; (22a) cannot survive.

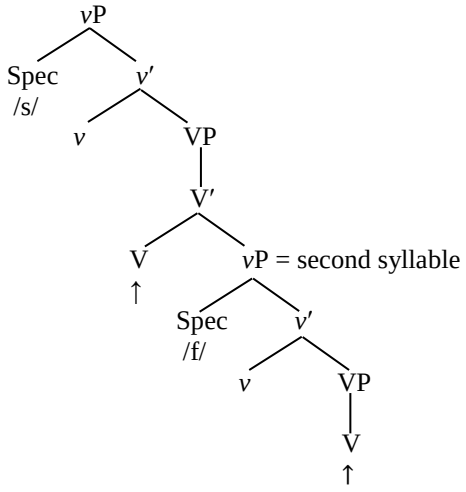
In a tense/lax system, in which spelling the Nucleus out at *v* or V is contrastive, a vowel can only be spelled out in *v* (i.e., ‘be lax’) if SpecVP is projected and occupied. This, we believe, is the quintessence of the *markedness* of lax vowels in languages such as Dutch: *(i)* the universal default is for the Nucleus to be spelled out at V (so tense vowels are inherently less marked than lax ones); and *(ii)* spelling the Nucleus out at *v* is allowed only if the *v* and V positions are separated by some non-vocalic melodic material associated with SpecVP, which is precisely what a lax vowel requires. Lax vowels are, in a sense, ‘obligatorily transitive’ (see Anderson 2011), like ‘affecting verbs’ – that is, they require a ‘theme argument’.

This simple approach to the distinction between tense and lax vowels, hinging on a difference in spell-out site (V versus *v*, resp.) which is afforded by the *v*-V approach, also gives us an account for the contrast between *sofa* (with tense ‘o’) and *koffie* ‘coffee’ (with lax ‘o’) (both with initial stress), the latter featuring what van der Hulst (1984, 1985, 2006) has called a ‘virtual geminate’. Let us start with the representation of *sofa*, which features a tense vowel in the first syllable. A tense vowel in Dutch is happy to occur in an ‘open’ syllable. In our terms, this translates into the statement that a tense vowel does not require filling of the Spec of VP – in fact, it cannot have SpecVP occupied. So the V-head of the first syllable is welcome to take as its complement the substring *fa*, represented as the second syllable in a trochaic foot[22] – i.e., a *v*P in the complement of the tense vowel, spelled out at V:[23]

[22] Indeed, our proposal that syllables can be embedded in syllables represents feet as such, rather than as sequences of syllables; see section 4 for further discussion.

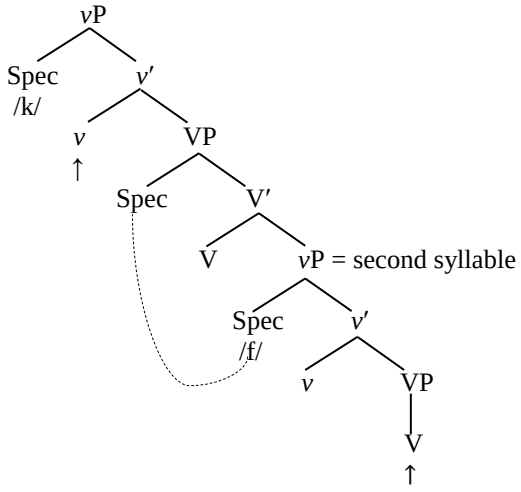
[23] We will see below that both a tense vowel and a stressed lax vowel plus following consonant (see *tempo*) can take *v*P as a complement.

(23) /sofa/ *sofa* ‘id.’



Now what happens when we are dealing with a lax vowel in the first syllable, as in *koffie* ‘coffee’? One thing that will change is the spell-out locus for the Nucleus: lax vowels are spelled out at *v*. But shifting the upward-pointing arrow from V to *v* in (23) brings about no substantive change: *v* and V are string-adjacent, so shifting the arrow from V to *v* is a vacuous operation when SpecVP is not projected (as in (23)). So in the representation of *koffie* ‘coffee’, SpecVP *must* be projected (as required by the ‘transitivity’ of the lax vowel), and it must in addition be associated with non-vocalic melodic content. In *koffie*, the SpecVP position can be occupied via base-generation (‘External Merge’ in current syntactic terminology) only by insertion of a sonorant (as in *tempo*). But there is no sonorant consonant in *koffie*, which means that the specifier position of VP is unoccupied in the base. If the SpecVP position remained unoccupied, the first syllable of this trochee could not contain a lax vowel. So SpecVP must get filled, but it cannot be filled here via External Merge. Thankfully, there is a way out of the dilemma: positions that are not filled via External Merge can be occupied in the course of the derivation via Internal Merge, i.e., the ‘recycling’ of material externally merged into the structure. So the dilemma posed by *koffie* is solved by ‘moving’ the /f/ into the SpecVP position, and making it simultaneously the Onset of the second syllable and part of the Coda of the first. ‘Movement’ should, of course, not be taken literally: the /f/ is not moving around the structure of the word; ‘movement’ is a metaphor. The way in which this metaphor has customarily been formally expressed in phonology is via *spreading*, or *multiple association*: the melodic material represented by /f/ is associated both with the Onset position of the second syllable and with the SpecVP position of the first:

(24) /kɔfi/ *koffie* ‘coffee’



In languages in which the doubly linked melodic material can be spelled out in both positions, this results in gemination. (Phrased in terms current in syntactic analysis, what we would say is that both copies in the chain are realized.) In Dutch (which does not have surface geminate consonants), the melodic material is spelled out just once, in the Onset position of the second syllable.[24] But importantly, this material is *also* associated with the SpecVP position in the first

[24] In *koffie*, the non-sonorant melodic material for the intervocalic consonant is externally merged in the Onset position (Spec*v*P) of the second syllable. It is spelled out there rather than in the SpecVP position of the first syllable, to which it spreads. A reviewer points out that in syntax, when a constituent externally merged in some relatively low structural position links up via Internal Merge to a position higher up the tree, it is usually the higher position that serves as the spell-out site. This is because this higher position is typically one in which some property important to the interface between syntax and semantics/information structure is satisfied. On the syntax/semantics side of the grammar, this is usually a cogent reason to spell the multiply associated element out in the position of Internal Merge. On the phonology side, other considerations play a role to adjudicate the locus of spell-out. Onset Maximization is one important such consideration. We submit that it is for this purpose that the multiply associated /f/ in (24) receives its surface exponent in the Spec*v*P position of the second syllable, not in the SpecVP of the first.

Note that Dutch orthography actually spells the ‘f’ twice. This convention is also used in the writing of words such as /kɔma/ *komma* ‘comma’ (cf. tense /koma/ *coma* ‘id.’) and /fɛlʉm/ *vellum* ‘id.’ (cf. tense /felʉm/ *velum* ‘id.’). For cases such as *komma* and *vellum*, in which the stressed lax vowel is followed by a sonorant consonant, it is sufficient for the licensing of vowel spell-out at *v* to have the sonorant associated just with SpecVP, where the sonorant is externally merged: spreading the melodic content of the sonorant down to the Onset position of the second syllable is not required for this purpose. Such spreading nonetheless does take place, however, with an eye toward satisfaction of Onset Maximization, which causes the intervocalic sonorant to be spelled out as the Onset of the second syllable. So in the representation *komma* and *vellum*, too, we postulate a link between the first syllable’s SpecVP position and the second syllable’s Spec*v*P; the difference with (24) is that while in (24) the melodic content is externally merged in Spec*v*P, in *komma* and *vellum* the intervocalic sonorant is externally merged in SpecVP and spreads to Spec*v*P. (Note that phrased in syntactic terms, this spreading is a case of ‘downward movement’, customarily thought to be prohibited in syntactic structures because the ‘trace’ left by such movement cannot be licensed. The adoption of the ‘copy theory of movement’ has made the apparent ban on downward movement much less obvious. This is an area where we hope that phonology can inform syntax. We plan to return to this in future research.)

syllable, which licenses the spell-out of the Nucleus at *v* thanks to the fact that the SpecVP position is projected and associated with melodic material (albeit covertly).[25]

At this point, it may be worth commenting in some more detail on the ways in which SpecVP can be used in the structure of the syllable, and drawing a useful parallel with syntax. In the preceding discussion, we had initially restricted SpecVP to *sonorant* consonants. But in the analysis of Dutch *koffie*, we allowed non-sonorant melodic material to 'spread' to SpecVP in a structure in which this position is projected but not filled via External Merge. Why doesn't the association of non-sonorant material with SpecVP violate the restriction on filling SpecVP exclusively with sonorant material? Larson's (1988) analysis of the double object construction suggests an answer to this question from a syntactic perspective. For Larson, the SpecVP position is 'ordinarily' the position into which the Theme argument is merged (as in *John gave a book to Mary*); but in the syntax of the double object construction (*John gave Mary a book*), Larson takes SpecVP to be occupied by the indirect object (i.e., the Goal, not the Theme; the latter is 'demoted' to adjunct status). The strongest possible interpretation of the Uniformity of Theta Assignment Hypothesis (UTAH; Baker 1988) would take it to establish biunique relations between particular thematic roles (here, the Theme) and structural positions (here, SpecVP) would lead one to expect that the SpecVP position ought to be uniquely and exclusively associated with the Theme role. But Larson is aware that Baker himself formulated the UTAH less strictly, in a way that leaves open precisely the kind of exploitation of SpecVP that Larson advocates. Baker's UTAH says that identical thematic relations between items should be represented by identical structural relations between those items at D-structure. For SpecVP in syntax, this means that whenever it is filled by an argument through External Merge (i.e., at D-structure), this argument will be a Theme; but if for whatever reason SpecVP is *not* filled via External Merge (e.g., because V is dethematized, as in Larson's analysis of the dative shift alternation), it will be free to be occupied by a non-Theme via Internal Merge. When we now return to SpecVP in our phonological representations, we see that it is subject to a restriction on External Merge that says that only sonorant consonants can be inserted there; but when SpecVP is structurally projected *without* being associated with melodic content through External Merge, it is free to be associated with non-sonorant melodic material via Internal Merge. The parallel with Larsonian syntactic structures is perfect. (By this, of course we do not mean to suggest that Larsonian syntactic structures *themselves* are perfect: we will not commit ourselves to any particular analysis of ditransitive sentences here.)

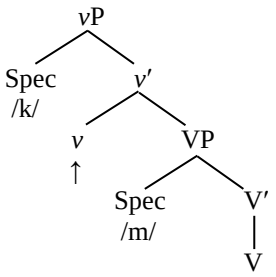
We have now derived an analysis of Dutch *koffie* and similar such disyllabic words with a lax vowel in the first syllable and a single non-sonorant consonant in intervocalic position which gives a particularly precise expression to van der Hulst's (1984, 1985) insight that the intervocalic consonant in such words is a 'virtual geminate'. The intervocalic /f/ in *koffie* is Externally Merged as the onset of the second syllable, but is also associated, via Internal Merge, with the SpecVP position of the first syllable. It is thanks to this association of /f/ to SpecVP that spelling out the Nucleus of the first syllable at *v* (i.e., realizing it as a lax vowel) is legitimate. We assume that if the single intervocalic consonant *is* a sonorant (as in *komma* 'comma') it is likewise externally merged in the Onset of the second syllable and internally merged into SpecVP, although in this the opposite could also be considered; in that case the

[25] Schwa is not restricted to occur only in a structural environment in which SpecVP is occupied. We assume that schwa is not the reflex of melodic content that is underlying present under *v* or V but instead is the surface realization of a Nucleus that is phonologically empty (i.e., not associated with melodic content).

Internal Merge is driven by the requirement that syllables prefer Onsets. However, we know of no argument to represent *koffie* as different from *komma*.

Next, let us consider how to analyze monosyllabic *kop* 'cup' and *kom* 'bowl', featuring a lax vowel followed in the first case by an obstruent and in the second by a sonorant. Here again, we need a license to spell the Nucleus out at *v*. Such spell-out is legitimate only if SpecVP is projected and associated with melodic material. For *kom*, this is easy to achieve:

(25) /kɔm/ *kom* 'bowl'



In (25), the nasal is Externally Merged in SpecVP, and the complement-of-V position is not used. This structure is well-formed as is: projection of SpecVP is not contingent on projection of a complement for V. The parallel with syntax is once again informative: in *There arrived a plane*, the notional subject is a Theme, which by Larson's (1988) and Hale & Keyser's (1993) application of UTAH must be base-merged in SpecVP; but the verb (*arrive*) here has no complement (unlike in *There arrived a plane at the airport*), so there is nothing sitting in the complement-of-V position.[26] In neither *There arrived a plane (at the airport)* nor *kom* 'bowl' is the V position radically empty: it is in a chain with *v*, which is where the head of the *v*-VP structure (the verb *arrive* or the lax vowel /ɔ/) is spelled out.[27]

Without the nasal in SpecVP, the structure in (25), with the arrow pointing to *v* as the spell-out site of the Nucleus, falls apart. Recall that with SpecVP unprojected, the distinction between *v* and V as the spell-out site for the Nucleus cannot be made. Shifting the arrow from V to *v* would be a vacuous operation; spell-out in *v* requires a consonant (via External or Internal merge) in SpecVP. It follows that without the nasal in SpecVP, (25) can only be realized as /ko/ (as in *Ko*, a proper name; with a tense /o/), not as */kɔ/. Put differently (but equivalently), when SpecVP is empty, the arrow can only point to V; and an arrow pointing to V delivers a tense vowel, in languages (such as Dutch, with its tense/lax distinction) in which the locus of vocalic spell-out is distinctive.

For *kop*, with a /p/ instead of a sonorant following the lax vowel /ɔ/, we cannot resort to External Merge in SpecVP: after all, /p/ is not a sonorant, so base-insertion of this consonant in SpecVP violates the phonological equivalent of UTAH (i.e., UMAH). But we *can* in principle

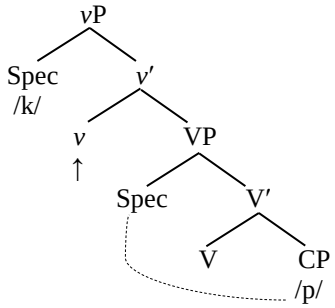
---

[26] In *He fell*, the Theme is base-merged in SpecVP but then raises to SpecIP for licensing purposes. In phonology, there is no movement into the specifier position of a functional projection, so the element base-generated in SpecVP stays put. This is precisely what happens in syntax in *There arrived a plane*, which is why we included this sentence in the main text alongside *He fell*.

[27] Here we are assuming that an unaccusative construction such as *There arrived a plane (at the airport)* contains a projection of *v*, despite the absence of an Agent. In approaches that tie the distribution of *v* to predication (such as den Dikken 2006), the presence of a projection of *v* in the structure of *There arrived a plane (at the airport)* can be straightforwardly ensured on the plausible hypothesis that *there* (the so-called 'expletive') is in a predication relation with the VP. We will not dwell on this matter further here.

insert /p/ in the complement-of-V position, and then associate its melodic content with the SpecVP position via 'spreading' (or 'Internal Merge'), as in (26). This creates a virtual geminate of sorts.[28] The fact that *kom* and *kop* end up with different structures finds some justification in Dutch, based on the allomorphy of the diminutive suffix: while *kom* 'bowl' forces schwa insertion (*kommetje* 'little bowl'), *kop* 'cup' does not (its diminutive is *kopje*, not *koppetje*[29]).

(26) /kɔp/ *kop* 'cup'



Note that in this analysis of the Dutch tense/lax distinction, no recourse needs to be had to a polysyllabic representation of an ostensibly monosyllabic word such as *kop* 'cup': the final /p/ in (26) occupies the Coda position of the single syllable constituted by *kop*; no second syllable with an empty Nucleus is necessary, unlike in Government Phonology approaches (Kaye, Lowenstamm and Vergnaud 1990; Lowenstamm 1996).[30] In van der Hulst (2010b), given that pairs like *sofa* /sofa/ and *sof* /sɔf/ 'bummer' are on a par in metrical terms, *sof* is represented disyllabically. In the current analysis, the equivalence is that a heavy syllable such as *sof* will initiate a 'foot' structure and be a 'foot' on its own, while *sofa* also constitute a single syllable which, as such, is a foot because /fa/ is an embedded syllable.[31]

---

[28] In this case, Dutch spelling does not use double consonants. A language like Swedish, where the situation may be similar phonologically, does use double spelling word-finally.

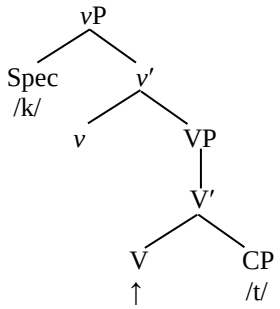
[29] We note that *koppetje* is not non-existent: it occurs as the diminutive of *kop* in its meaning of 'head'. For *pop* 'puppet, doll', the diminutive with schwa insertion (*poppetje*) also occurs alongside *popje*; see van der Hulst (2006) for detailed discussion of the Dutch diminutive.

[30] In Government Phonology's 'strict CV' model, Vs (and Cs) govern and license other Cs and Vs, but they are not joined in tree structures. Takahashi (1993) argues that positing both structures and government relations introduces a redundancy. But in syntax at least, government relations (now called Agree relations) are *defined* in terms of structure: c-command is a prerequisite for government/Agree, and c-command is a relation between nodes in a tree. The c-command relation is indispensable in the account of non-local dependencies. If all dependencies in linguistic structures were spec-head and head-complement relations, relations would be superfluous. But both in syntax and in phonology, dependencies/relations seem to be able to reach beyond the spec-head and head-complement configurations. It may be possible to recast apparently non-local relations in a local way; but that is not something this chapter can meaningfully address. Unless and until this recasting is successful, it seems to us that relations (in particular, c-command/Agree) remain necessary; and structures certainly are, too.

[31] If it should turn out, after all, to be essential for metrical reasons to represent *kop* and *sof* as disyllables, this can be achieved in our proposal by drawing yet another parallel with syntax, this time in the realm of 'object shift' and 'exceptional Case-marking (ECM)'. In Bošković (1997, 2002), it is argued for English that 'object shift', which we can represent as movement of a DP to SpecVP, is merely optional for direct objects of verbs (as in *John admires Bill*) but obligatory in the case of overt subjects of non-finite (small) clausal complements (as in *John considers Bill (to be) a genius*). Bošković's proposal thus makes a distinction with respect to association with SpecVP between the complement of the verb, on the one hand, and the specifier of the complement of the verb, on the other. Suppose that we carry this distinction over to syllabic structure, and differentiate with respect to association with SpecVP between the complement of the V-head, on the one hand, and the Onset of the syllable in the

The active ingredient in the analysis throughout is the distinction between spell-out at *v* (for lax vowels) and spell-out at V (for tense vowels), in conjunction with a particular hypothesis regarding the licensing of 'Nuclear' spell-out at *v*. In this theory, the difference between lax /kɔt/ *kot* 'cot' and tense (but not long) /kot/ *koot* 'talus' is made very straightforwardly with reference to the locus of spell-out of the Nucleus: a lax Nucleus is spelled out at *v*, and requires SpecVP to be projected and associated with melodic material; a tense Nucleus is spelled out at V (the default spell-out site for syllable Nuclei), and cannot have SpecVP projected (see (27) for the structures with tense vowels). We can think of the inverse correlation between occupancy of SpecVP and occupancy of V as a kind of 'doubly-filled Comp effect' familiar from syntax (yet not applicable, in syntactic representations, to the VP configuration): when the Nucleus is spelled out at V, its specifier cannot occupied by an element with independent melodic content.

(27) /kot/ *koot* 'talus'

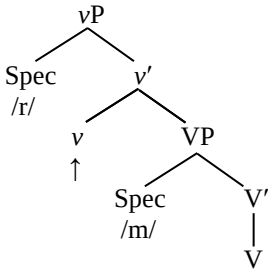


The syllable in (27) is called a 'superheavy syllable'. Such a syllable can end in an obstruent or in a sonorant (see (28c)). Superheavy syllables can either have a tense vowels followed by a tautosylalbic consonant, as in (27) and (28c), or lax vowels followed by two tautosyllabic consonants, as in (28d). In (28) we compare the structures of heavy and super-heavy syllables. What structurally distinguishes the heavy syllables in (28a) and (28b) from the superheavy ones in (28c) and (28d) is that in the latter, two positions in VP are associated with melodic content via External Merge: in (28c) both SpecVP and the complement-of-V position are occupied, and in (28d) melodic content is base-generated in the V and the complement-of-V positions; by contrast, in (28a) and (28b) only one position in the VP (SpecVP in the former, and the complement-of-V position in the latter) is filled via External Merge.

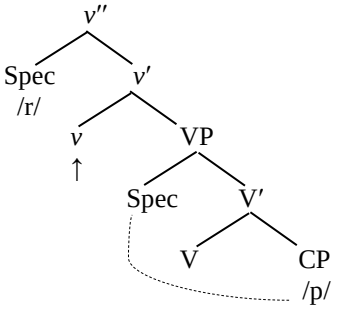
---

complement of the V-head of the first syllable, on the other. (Note that the subject of the non-finite clausal complement of a verb is in a geometrical relation with the matrix verb that is entirely on a par with the geometrical relation between the Onset of the second syllable of a trochee and the V-head of the first syllable: in both cases, the relation between the matrix V and the specifier is a 'niecehood' relationship.) If in addition we strengthen the distinction into a genuine dichotomy, we arrive at the result that association of a postnuclear non-sonorant consonant with the *specifier* of VP is possible only if it is mapped into the Onset (i.e., the *specifier*) position of a second syllable, in a trochaic foot whose second Nucleus remains unpronounced because it is properly governed (in the sense of Government Phonology) by the V-head of the first syllable. It is thanks to its occupancy of the *specifier* position of the second (silent-headed) syllable that /p/ has the license to 'spread' to the *specifier* position of the VP in the first syllable. On this approach, *kop* 'cup' is like *koffie* 'coffee' with respect to the structural position of the obstruent, with the difference between the two being that there is no overt second verb 'at the bottom' in the former.

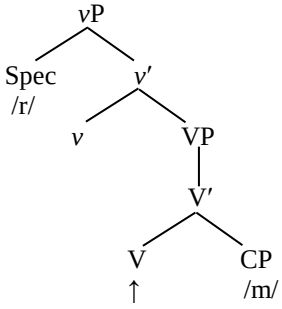
(28) a. /ɾam/ *ram* ‘ram’ (cf. (25))



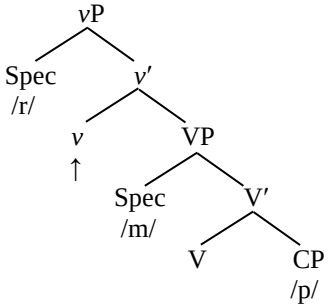
b. /ɾap/ *rap* ‘fast’ (cf. (26))



c. /ram/ *raam* ‘window’ (cf. (27))[32]



d. /ɾamp/ *ramp* ‘disaster’



The structure for *ramp* differs from the ones for *ram* and *rap* in being ‘ditransitive’: both dependent positions in the Rhyme are occupied via External Merge, one (SpecVP) by the nasal and the other (the complement-of-V position) by the stop. Both *ram* and *rap* are ‘monotransitive’, but, as we have seen, in different ways: External Merge here targets only one dependent position in the Rhyme, but in *ram* the dependent is in SpecVP whereas in *rap* the dependent is externally merged in the complement-of-V position, and ‘spreads’ to SpecVP via Internal Merge.[33] The External Merge sites of the closing consonants in *ram* and *rap* are different, but what unites the two cases and distinguishes them as a pair from *ramp* is that they both have just a single consonant in the Rhyme. ‘Ditransitive’ *ramp*, by contrast, has both

[32] Phonetically, as Gussenhoven (2008) has shown, the tense vowel in *raam* is truly long. We do not represent this structurally: the /a/ occupies just a single V-position in the structure. As the text discussion above (28) shows, the fact that *raam*, like *koot* (where the /o/ is not phonetically long), is superheavy follows without the /a/ being assigned two spots in the structure.

[33] Along these lines, our structures make the distinction between a monomoraic syllable (cf. the structure in (22b), for *thee* ‘tea’), a bimoraic one (*kop*, *kom*), and a trimoraic one (*ramp*) in terms of the vowel (one mora) plus the number of internal ‘arguments’ (with each adding a mora).

SpecVP and the complement-of-V position occupied via External Merge, making it 'superheavy'.

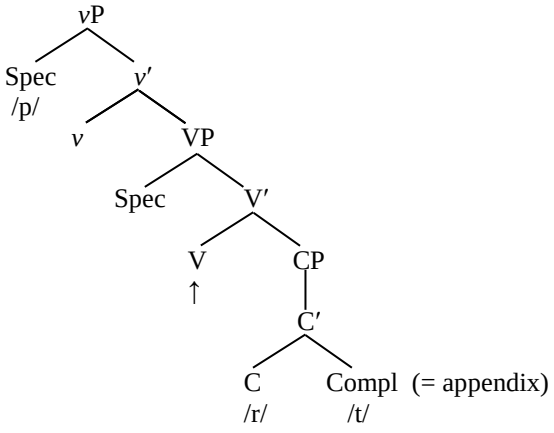
3.5 A note on Coda clusters

In Dutch /ram/ *raam* 'window', with the tense vowel /a/ spelled out in V and SpecVP being unfillable when V is occupied, the complement-of-V position is the *only* position in which /m/ can be inserted when it follows a tense vowel.[34] This leads us to predict that it should not be possible to fill the complement-of-V position with some other consonant *in addition to* the sonorant. Likewise, a syllable like *ramp* can also not be augmented with another consonant. This prediction is borne out by the fact that */ramp/ **raamp* and */rɑmpk/ **rampk* are impossible.

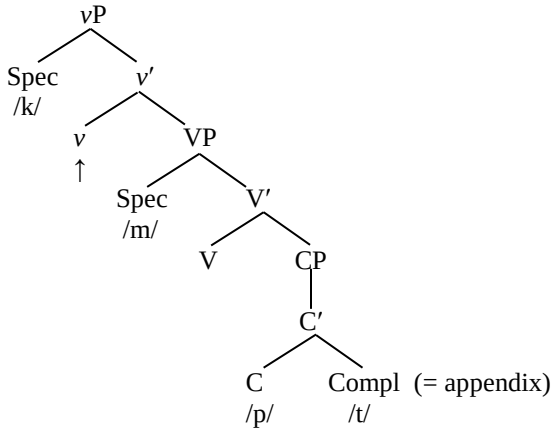
However, it would appear that superheavy syllables can be augmented, but only if the extra consonant is a coronal:[35] /mant/ *maand* 'month' is grammatical alongside /man/ *maan* 'moon', /start/ *staart* 'tail' is grammatical alongside /star/ *staar* 'cataract', and /falt/ *vaalt* 'dung heap' occurs alongside /fal/ *vaal* 'faint, pale'; all these words have the same tense /a/. Likewise, /rɑmpt/ is a possible sequence, although, as it happens, only as an inflected form, as in the 3rd person singular of the verb *kampen* 'struggle': *het bedrijf kampt met grote verliezen* 'the company is struggling with large deficits'.

To accommodate these cases, we exploit internal complexity inside the Coda constituent:

(30) a. /part/ *paard* 'horse'



b. /kɑmpt/ *kampt* 'is struggling'



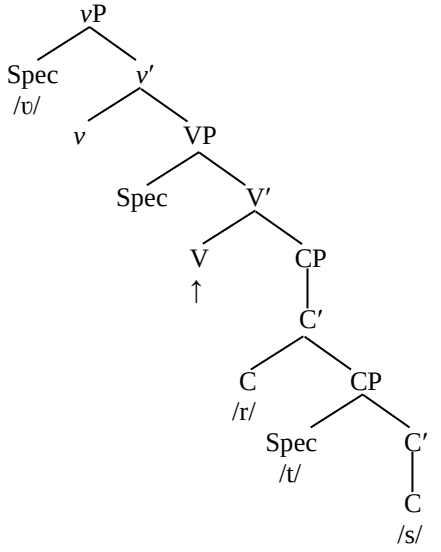
---

[34] Note that the sonorant consonant in *raam* 'window' is inserted not in SpecVP but in the complement-of-V position, and that it does not get associated with SpecVP at all because the tenseness of the vowel allows the SpecVP position to remain entirely unprojected. Does this contradict what we had said previously about the locus of sonorants? It does not: the relation between specific melodic content and specific structural positions is not biunique. What UMAH in (15) says, for SpecVP, is that it can uniquely be base-filled by sonorants – put differently, sonorants are the only segments that can be base-inserted in that position. This is crucially not the same as saying that sonorants can only be inserted in SpecVP. Sonorants can show up in *any* position in the syllabic template (recall also the brief discussion of *komma*), even including the Nucleus position (in the case of syllabic nasals and liquids): they are truly *factotum* elements. So there is no problem with the fact that /m/ is in the complement-of-V position in (28c): the complement-of-V position is a perfectly legitimate position for sonorant.

[35] From a synchronic point of view, there is no obvious explanation for this restriction to coronal 'augments' other than an appeal to the widely acknowledged 'unmarked' status of coronal place. See also section 5.5 for relevant discussion.

The Coda cluster can actually be made even more complex by adding a /s/ after the /t/, as in the second syllable of *voorwaarts* ‘forward’, transcribed as /ʋarts/. We can accommodate /t+s/ in the complement of the liquid.

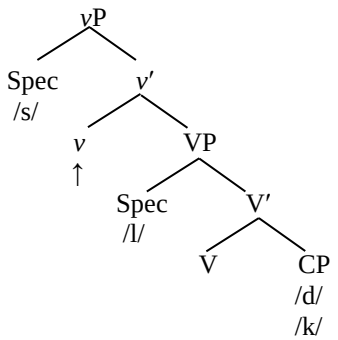
(31) /ʋarts/ *(voor)waarts* ‘(for)ward’



The restriction that a superheavy syllable in Dutch can be followed by a consonant cluster only if the extra consonant(s) is/are coronal has an interesting companion in English. There, when the syllable Nucleus is a tense vowel, a sonorant consonant that follows it can, in turn, be followed by another consonant only if this is a coronal (thus, *wield* but not **wielk*, and *(Glenn) Gould* but not **goulp*); by contrast, a lax vowel + sonorant sequence can readily be followed by a non-coronal consonant (so that alongside *silt* we also find *silk*). The proposal for Dutch straightforwardly extends to these English cases.

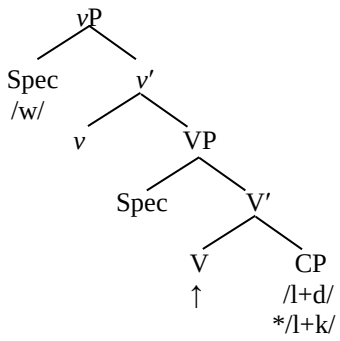
Following a short/lax vowel that is spelled out in *v*, a postvocalic sonorant+stop sequence can be accommodated in the structure of the syllable by mapping the sonorant into SpecVP (recall that sonorants have the unique license to be base-generated in SpecVP), so that the stop following it can occupy the complement-of-V position all by itself, as in (32) – the representation of English *silt* and *silk*.

(32)



But a tense vowel, spelled out in V, precludes occupancy of SpecVP. The postvocalic sonorant must be mapped into the complement-of-V position. If followed by a stop that is also mapped into the Coda, the sonorant must form a cluster with the stop in the complement-of-V position. Thus, for English *wield*, we arrive at (33) as the structure of the syllable.

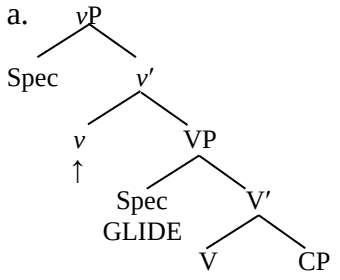
(33)



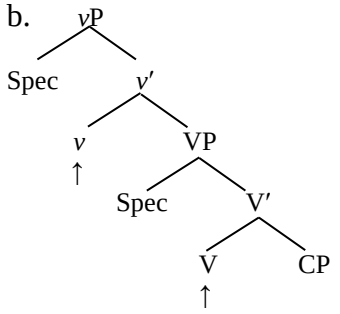
Since the sonorant in SpecVP in (32) does not form a consonant cluster together with the stop in the complement-of-V position, the two can be specified for place information entirely independently of one another, and there is no requirement that the stop be coronal. This freedom is absent in the presence of a tense vowel because, with the vowel spelled out in V, SpecVP is unavailable for External Merge of the sonorant; this consonant must hence be mapped into the complement-of-V position and form a consonant cluster with the following stop, and (in English just as in Dutch) such a cluster is well-formed only if the stop is coronal.

The restriction on liquid+stop sequences also applies in the case of long lax vowels and diphthongs (e.g., *mold* but not **molk*). We can understand this when we examine the representation of these Nuclei in English. Following Szigetvári (2016), we represent English long lax vowels such as /ɔ:/ and /ɑ:/ as in (34a), with /h/ as a glottal glide occupying SpecVP;. This representation carries over to the diphthongs /ɑj/, /ɑw/ and /ow/ as well. And (34b) is a natural representation in our system for the diphthong /ɪə/ in *weird*.

(34) a.



b.



In both (34a) and (34b), the SpecVP position is unavailable for a consonant that is not part of the long vowel or diphthong – in (34a) because the position is occupied, and in (34b), just as in the case of tense vowels, because the V-head is filled (which precludes occupancy of SpecVP).[36] So here again, although sonorant consonants can in principle be merged in SpecVP, this opportunity is blocked. The postvocalic sonorant must therefore be mapped into the complement-of-V position, just as in (33).[37]

4 Foot structure in X-bar phonology

In section 2 we discussed how the notion of embedding one syllable inside another entails a different perspective on foot structure: a bisyllabic (trochaic) foot is replaced by a structure (which can as such form a foot, given that foot structure is independently needed; see below) in which one syllable is embedded in another (recall (6)). In the present section, we would like to make some tentative proposals for how feet might be (re)analyzable in a theory that acknowledges syllable embedding. Looking at the question from a general, *a priori* perspective in the context of our 'phonology is syntax' program, there is a first division to be made between *(a)* foot structures in which one syllable is embedded within another in such a way that it occupies a structural position made available in the *v*-V structure of the syllable, and *(b)* foot structures in which there is no such embedding. For examples of type *(a)*, in sections 2 and 3 we have already come across the plausible case of a subordination approach to trochaic feet, with the subordinate syllable in *complement* position. Depending on one's theory of foot structure, this is where the reanalysis of foot structure in terms of recursive syllable structure could stop. Various students of stress have argued *against* any other foot type, including iambic feet, either only when weight-sensitive (Hayes 1995; Kager 1993) or more generally (cf. van de Vijver 1998 and van der Hulst 1997). Whatever the merit of these proposals, we will here explore what kind of structures might be entertained to capture prosodic 'WS' units.

Formally speaking, bearing in mind syntactic analogues, reanalyses of foot structure, as well as of 'higher' prosodic units, could involve embedding, adjunction or coordination. In the following subsections, we will first repeat out proposal for capturing trochaic feet in terms of embedding 'syllables inside syllables', adding that an apparent trochaic unit might also result from adjunction. We will then turn to iambic patterns, proposing to analyze these as 'derived structures'. Finally, we consider the issue of prosodic (or phonological words), which we propose to analyze in terms of coordination.

4.1 Trochees

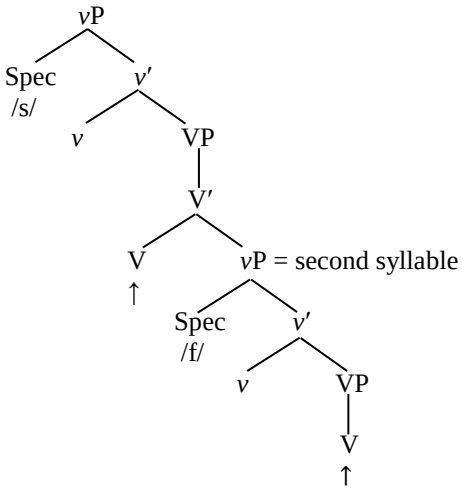
In the discussion in section 3, we discovered that a syllable (i.e., *v*P) can occur as the complement of the V-head of the preceding syllable. When this occurs, the structure that is derived captures the idea of a trochaic foot: the second syllable is structurally subordinate to the first

[36] Moreover, filling SpecVP with consonantal material in (34b) would cause the exponents of *v* and V to become discontinuous, hence unpronounceable as a diphthong.

[37] A reviewer asks what our account of the Rhyme of *excerpt* is. In the rhotic pronunciation of this word, a short schwa (spelled out in *v*) is followed by /r/ in SpecVP, and a legitimate /pt/ Coda cluster (cf. *apt*, which has the same gross structure). Here we are not dealing with a tense vowel, long lax vowel or diphthong followed by a sonorant plus obstruent cluster. We regard the non-rhotic version of *excerpt* as a phonetic variant of the rhotic one.

one, and located on the recursive side in a right-branching structure, with the Nucleus of the first syllable as the head of the structure. To illustrate, let us repeat the structure assigned to Dutch *sofa* in section 3.4 as in (35). Since the Nucleus of the first syllable is the head, it attracts the stress, resulting in the strong-weak pattern defining the trochee.

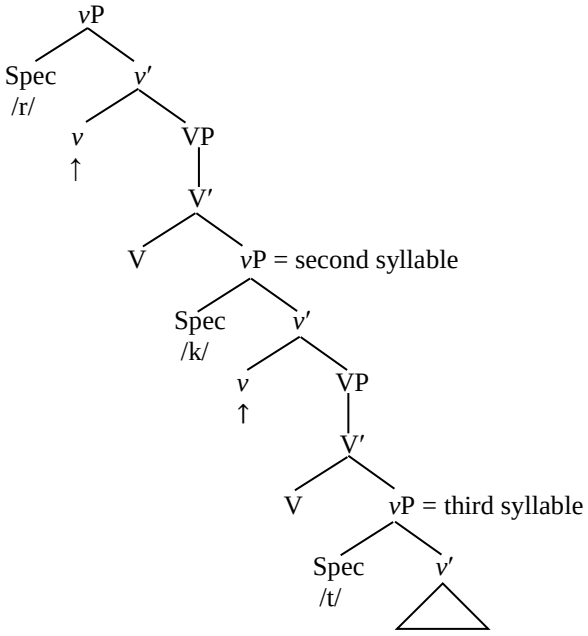
(35) /sofa/ *sofa* 'id.'



For dactyls (i.e., feet with the stress pattern σ́σσ, where the underscore marks stress, such as *rickety* and *vanity*), for which traditional metrical phonology requires a ternary foot, the syllable subordination approach that we are advocating for trochees makes a simple extension available involving two levels of embedding, as in the structure in (39).[38] Van der Hulst (2012) draws attention to the fact that the structures in (35) and (36) provide a structural basis for poetic rhyming patterns that involve these entire structures minus the highest onset. Traditional foot structure provides no such account.

[38] In the structure in (36), we suppressed the multiple association of the stops /k/ and /t/ with the SpecVP positions above them. Recall the discussion in section 3.4 of the fact that a tense/lax distinction can be made only in structures in which *v* and V are separated by melodic material that is not associated with *v* or V (or, more simply put, in structures in which SpecVP is projected). For the discussion of foot structure in this section, this detail is immaterial.

(36) /rɪkɪti/ *rickety*



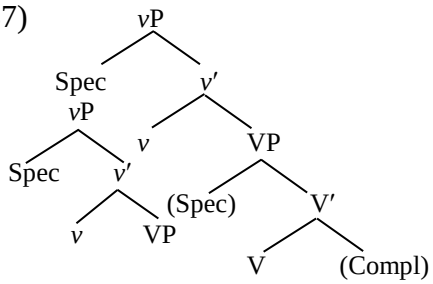
An obvious objection to this proposal is that, in principle, we could represent more complicated structures with additional degrees of embedding. But the naked fact that infinite embedding is a formal-theoretical option does not imply that natural languages impose no limits on such embedding. In syntactic structures, processing considerations curtail multiple embedding (van der Hulst 2010a). Likewise, processing considerations of a different kind limit recursion in phonology. In section 6 we discuss briefly which 'forces' are at work in phonology to make structures that go beyond two degrees of embedding unlikely.

It is generally the case, in syntax, that recursive structures can either result from complementation or from adjunction. There is no *a priori* reason to reject the same two options in phonology. The trochaic structure is recursive because the complement (i.e. dependent) of a head is identical to the maximal projection of the head. This is the kind of recursion that is illustrated in (36). The dependent, which causes recursion, is a complement to the head. But a V-headed structure (a syllable) should in principle also be embeddable inside a larger V-headed structure as an adjunct. Adjunction of a syllable to a trochaic foot would deliver a 'superfoot', which has been the usual account of dactylic patterns. However, if indeed such SWW patterns are structurally ambiguous (resulting from embedding or adjunction), it behooves us to ask whether the occurrence of one or the other can be positively identified. Presumably, as a general rule of thumb, adjunction is resorted to as a last resort: the structure-building engine's first resort will always be to exploit complementation and specification. We see this in syntax, too. What this could mean in phonology is that the adjunction structures result from syllables that remain unparsed, especially in weight-sensitive systems, because they are simple too small (too light) to form a foot by themselves.

4.2 Toward a representation for iambic feet

Moving on to iambs (WS patterns), consider first the structural option of having the specifier position of the *v*P occupied by an entire syllable – i.e., by another *v*P, as in (37):

(37)



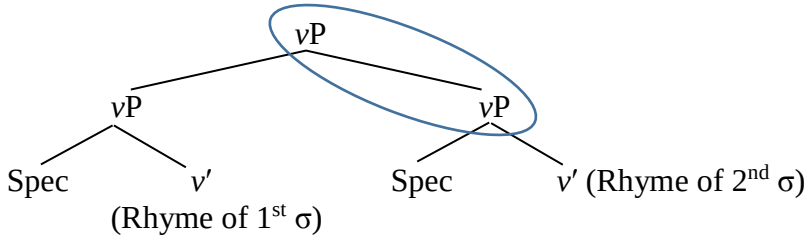
At best, (37) could only represent iambic feet whose stressed syllable lacks an Onset: after all, the first, unaccented syllable occupies the specifier position of the stressed syllable; the specifier position of *v*P is ordinarily the position for Onsets. To the extent that iambic feet exist in which the accented syllable is (necessarily) Onsetless, they might be candidates for the structure in (37). But no iambic foot whose second syllable has a true Onset could ever be represented in these terms.

While this problem suffices to reject (37) as a structural option for iambs on phonological grounds, there is a further general consideration – emerging from our research program to establish structural analogies between phonology and syntax – which could potentially explain why there are no iambic feet structured as in (37). From syntax, we are familiar, from a variety of different contexts, with the apparent fact that 'bare' propositions (small clauses and complementizerless tensed clauses) very strongly tend not to occur as subjects of predication. As an illustration, consider the following. The sentences *I saw John leave*, *I saw it happen*, and *It happened that John left* are all fine. But 'squeezing' the first two sentences into one by replacing *it* with [*John leave*], which would be semantically perfectly coherent, delivers an ill-formed result: **I saw John leave happen*. By contrast, the *it* of *I saw it happen* can readily be associated with a proposition in 'extraposed' position (as in the third sentence), yielding *I saw it happen that John left*. The ungrammaticality of **I saw John leave happen* is directly germane to the question of whether (37) could represent a well-formed (iambic) foot. In **I consider* [[*John leave*] *happen*], we have one verbal small clause (*v*P) embedded in another, as its subject/specifier. The result is woeful (regardless, in fact, of the category of the small clauses: **I consider* [[*John smart*] *obvious*] is also impossible). As a general rule, 'bare' propositions (subject-predicate structures) cannot be embedded as specifiers inside larger propositional structures. A variety of attempts have been made in the syntax literature to understand this ban. But since we are not aware of an explanatory proposal that covers the entire range of cases, we will content ourselves here with stating what appears to be an empirical generalization: in syntax, 'bare' propositions (subject-predicate structures) cannot be embedded as specifiers of propositions.

Of course it could be that the root of this generalization lies in the semantics – the fact that we have phrased it in terms of propositions (a semantic notion) may be indicative of this. If so, this generalization may not tell us anything about whether (37) is or is not legitimate in metrical phonology. But we actually suspect that we are dealing here with a deeply structural restriction on specification structures, and will henceforth consider (37) not to be grammatical.

If, then, the structure in (37) is not an option for iambic feet, what to do with such feet, if they truly exist? One intuitively highly plausible way to model the structure of iambic feet in line with the syntax-inspired X-bar-theoretic approach is to treat the first syllable of an iambic foot, on the analogy of syntax, as a *TOPIC* (as in *Mary, I really like*) rather than as a subject:

(38)



In the structure in (38), the ellipse highlights the 'host' structure. Repetition of the *v*P level of the host is indicative of adjunction, as distinct from specification.[39] This structurally marks the initial unstressed syllable of an iambic foot as *extraprosodic*, in the same way that the initial topic of a topicalization construction in syntax is structurally marked as *extrasentential* (and usually not pitch-accented).

The unaccented syllable of an iambic foot is welcome to have both an Onset and a Coda, as is the stressed syllable: their internal structures are in fact entirely independent of one another. We thus never expect Codas to 'shift' over to Onset positions, nor is ambisyllabicity expected in iambic feet. This is all as it should be.[40]

The syntax literature contains many examples of topicalization phenomena featuring so-called 'connectivity effects' – effects which suggest that the topic binds a copy in clause-internal position that remains either entirely silent (as in the case of 'ordinary' topicalization: *Mary, I don't like*) or partially silent (as in analyses of certain left-dislocation phenomena; cf. German *Maria, ich mag <u>die</u> nicht*, where *die* is a resumptive pronoun in clause-internal position). We could now imagine that iambic feet of the type in (38) would also have the *v*P-adjoined syllable in initial position bind a (partially) silent copy in the complement of the V-head of the accented syllable, and could show 'connectivity effects' (harmony) via this copy. On an analysis of this sort, such iambs really are not underlying feet at all: what underlies them is a trochee whose second syllable is silenced (in part or in full). Along this path, we get a novel and productive purchase on van de Vijver's (1998) conclusion that only trochees exist and that iambs should be represented in trochaic terms.

---

[39] A reviewer finds that such adjunction has no analogue in syntax. We disagree. The existence of 'scrambling' phenomena is indicative of the existence of topicalization at the level of the 'bare' predication, *v*P (see also the cartography literature (e.g., Belletti 2004) on low topic positions – customarily represented there with the aid of designated functional projections; we reject functional projections for phonology, and consider there to be underwhelming support for TopP-structures in syntax). Admittedly, '*Mad Magazine* sentences' (*Me wear a tie?!*), which may be the closest thing in syntax to a *v*P occurring by itself as the root, do not allow topicalization (*A tie, I would never wear* versus **A tie, me wear?!*), but this likely has to do with the speech act involved (cf. rhetorical yes/no-questions, which likewise resist topicalization: **A tie, would I ever wear?!*).

[40] The initial syllable of an iambic foot, occupying a position outside the X-bar structural core (a topic or left-dislocated position), is not in a properly governed position: we know this from syntax, where we can point to the impossibility of complementizer omission as an indication to this effect (cf. *Every sane person knows (that) Trump is a buffoon* vs. **(That) Trump is a buffoon, every sane person knows*). So the nucleus of the unstressed initial syllable of an iamb is not licensed to remain silent. This probably explains the English pronunciation of *Cnute*, with its intrusive schwa: in the iambic foot with /kV/ as the adjoined initial syllable, the Nucleus cannot remain silent because it is not properly governed; a schwa obligatorily spells out the Nucleus. The fact that the other Germanic languages pronounce this name with an initial /kn/ sequence indicates that in these languages this sequence can be represented as a legitimate Onset cluster whereas in English it cannot be. In indigenous words which historically have /kn/ Onset clusters, such as *knee*, English has 'solved' the problem not by constructing an iambic foot with an intrusive schwa as the Nucleus of the adjoined initial syllable, but by cluster reduction (via deletion of the /k/).

Given that iambic feet are always weight-sensitive (Hayes 1995), consisting of a heavy, 'bimoraic' stressed syllable that is preceded by a light unstressed syllable, we propose that the structure in (38) results from adjoining a syllable to a structure that it itself a (monosyllabic) *trochaic* foot. As per a proposal in van der Hulst and Ritter (1998), who provide an analysis of so-called minor syllables in Kammu, the resulting structure could, in fact, be called 'prosodic word'.[41] Adjunction of a weak syllable to a following trochaic foot is independently required for initial unstressed syllables in languages like English (as in *balloon* or *rebellion*; the second example, where the weak syllable is adjoined to a 'bisyllabic' trochee produces an amphibrachic (WSW) structure).[42]

4.3 Coordination

While we expect that adjoined syllables are weak, unstressed (or even lacking an overt vowel, as do minor syllables in Kammu)[43], forming units that are usually called (weight-sensitive) iambic feet, colon or even prosodic words, we need to also address the question how full-fledged feet combine into prosodic words that account for primary stress, i.e. represent which foot is the head foot of the word.

We here assume that the subordinate, embedded constituent in the structure of a foot is by definition the weak member and as such intrinsically light (just like adjoined syllables). A heavy syllable thus cannot be embedded. So in a SW Dutch word like *súltan*, *sul* and *tan* (both heavy) will have to form a structure different from any of the ones considered so far in this section: complementation and adjunction (which can only accommodate light dependent syllables) are both unsuitable.[44]

In syntax, there is one more relationship, besides complementation, specification and adjunction, which two constituents can be engaged in: the coordination relation, seen in conjunction and disjunction constructions (*John and/or Bill*). Coordination used to be represented in terms of ternary (or n-ary) branching, with the con/disjunction particle and the con/disjuncts grouped together into a flat structure. But more recent work in syntax has discovered that coordination obeys the binary branching hypothesis. One argument for this is the fact that *Every man and his wife came to the party* allows for a bound-variable interpretation of *his*, whereas *His wife and every man came to the party* does not – something that follows if the first conjunct asymmetrically c-commands the second: the first example is then a garden-variety case of bound-variable pronoun binding, and the second can be assimilated to the 'weak crossover' effect seen in *His wife loves every man*, which likewise makes no bound reading for

---

41 As suggested in van der Hulst (2000: 120), it is possible that a language with a prosodic 'colon' unit (Hayes 1995:217) also displays multiple right-strong prosodic *words* within the domain of a (long) morpho-syntactic word.

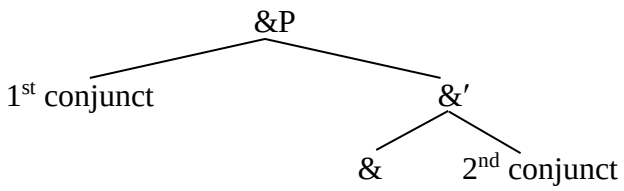
42 We refer to Martínez-Paricio (2013) for a general theory of 'layered (recursive) feet' which, in our approach, all involve adjunction.

43 This may be a modality effect with no counterpart in syntax: phonological recursion in our view encodes rhythmic structure (see section 6, where we make this explicit). Thus, if a closed syllable is always heavy, it will initiate (i.e., be modeled as) a matrix syllable.

44 The conclusion that *sultan* does not have a recursive foot structure is supported by the fact that the diminutive of Dutch *sultan* is *sultannetje*, with schwa insertion, and not **sultantje*: the second syllable (*tan*) behaves like a stressed syllable despite being less prominent than *sul*. There are some cases like this for which some speakers accept the short form of the diminutive; for such cases one might want to postulate a metrical representation involving complementation (as a true trochee), with the second syllable 'reduced' in some way; see van der Hulst (2008).

his available. Facts of this sort favor an analysis of coordination in which the first conjunct serves as the specifier of a phrase that contains the conjunction and the second conjunct:

(39)



This facilitates an analysis of feet of the *sultan* type, where each of the two constituent syllables is closed and heavy, and the relationship between the two must be such that neither is linked to the other as a dependent, via complementation or adjunction.[45]

The head '&' usually has an overt lexicalization in simple two-way coordinations in languages such as English: *John and/or Bill*. So-called asyndetic coordination (with a silent '&') is possible cross-linguistically for such simple coordination constructions; and in coordination constructions with more than two con/disjuncts, one often finds that all but the last &-head remain silent (*Tom, Dick and Harry*).[46] The fact that in syntactic coordination the deepest conjunct pair behaves differently from preceding conjuncts (in being more likely to take a non-silent conjunction) may have an interesting parallel in phonology. It has been shown that the deepest pair of feet, in a right-branching structure, may behave differently from higher structure, as captured in the occurrence of a SW relationship for the deepest foot pair. See van der Hulst (1984) for an analysis of Dutch stress which states that in the phonological word the right conjunct is labelled strong if and only if it branches.

In conclusion, we propose that 'feet' are combined into the phonological word via conjunction.[47] If the rightmost, structurally deepest foot carries primary stress this means that the word tree has a right-branching structure, which branching nodes being labelled as 'strong'.

5 X-bar structure inside segments and segmental integrity

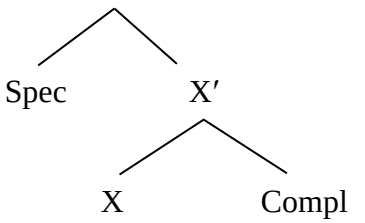
In this section, we extend the X-bar-theoretic approach to phonological structure to the internal structure of segments, representing the segment as an X-bar projection of a manner component, with laryngeal and place specifications accommodated in the specifier and complement positions, resp., of this X-bar structure. That is, we now delve into a development of the structure in (4b), repeated here, in pursuit of the hypothesis that the fact that segments have an X-bar-theoretic organization of the type in (4b) prevents them from taking additional

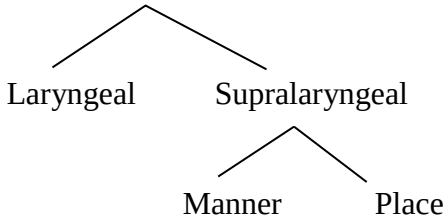
[45] In coordination structures in syntax (at least in Germanic), prosodically the most prominent member of the structure is usually the last one (cf. *Tom, Dick and HARry*). In *sultan*, for which the main text suggests a coordination approach, primary stress falls on *sul*, the first conjunct. This is not necessarily a contradiction: stress rules work differently at different levels. But this is certainly a matter that should be looked into further if the coordination approach to words like *sultan* is to be successfully pursued.

[46] The circumstances under which the &-head can or must remain silent need not concern us here: what matters is that a silent allomorph of & exists.

[47] Of course, we also need to look at the alternative of invoking adjunction. It has been argued in van der Hulst (1996, 2012) that the assignment of primary stress take priority over secondary, rhythmic stress. This means that the foot that expresses primary stress is assigned first. The subordinate status of other feet can then follow from recursively adjoining feet to the primary stress foot.

complements or specifiers outside this structure. We will examine what kinds of consequences this has for the relationship between segmental and suprasegmental phonology.[48]

(4) a. 

b. 

5.1 Vowels as syllable heads

One immediate implication of pursuing X-bar-theoretic approaches to both the segment and the syllable lies in the delimitation of the candidate set for the function of syllable Nucleus. It is of course perfectly well known that Nuclei are usually vowels, though syllabic consonants also exist. It is not as well known, perhaps, that what unites the kinds of consonants that can be used as syllable Nuclei with the set of vowels is the fact that these are all lack a contrastive specification for laryngeal properties (voicing, in particular[49]). Vowels are voiced by default (though voiceless vowels have been reported to exist, e.g. in Japanese, here voicing is not contrastive); and syllabic consonants are typically sonorants (liquids, nasals), for which voicing is also not distinctive: liquids and nasals can be devoiced, but this is usually an effect of their environment, such as the devoicing of liquids following stops in the Onset position of a stressed syllable in English; syllabically used sonorants are never contrastive for laryngeal properties. Why should there be this correlation between being usable as a syllable Nucleus and lacking a distinctive specification for laryngeal properties?

The answer to this question is straightforward, given our X-bar-theoretic outlook on the structure of the segment and the structure of the syllable. Syllables are *v*Ps, with the Onset as the specifier and the Coda (if present) in the VP (either in SpecVP or in the complement-of-V position). Syllables typically, perhaps invariably, have an Onset (with /ʔ/ as the Onset of apparently Onsetless syllables). Their Spec*v*P position is occupied by this Onset. Laryngeal information, whenever distinctive, is also projected as a specifier: recall (4b). If, as standard X-bar theory has it, there is exactly one specifier per head, it follows that the nucleus, whose

[48] In Mutlu (2017), very intelligent use is made of X-bar structure 'below the head', in the representation of the internal structure of segments and also in the representation of the structures resulting from the combination of segments (syllables). Throughout her work, Mutlu exploits the complement-of and specifier-of relations to great effect. It seems to us, however, that she goes too far in this exercise, by allowing constructs that themselves already have a specifier and/or a complement to in turn take a specifier and/or a complement higher up the tree. The most restrictive X-bar-theoretic hypothesis is to say that the fact that segments have an X-bar-theoretic organization of the type in (4b) *prevents them from taking additional complements or specifiers* outside this structure. This is the premise of the discussion to follow.

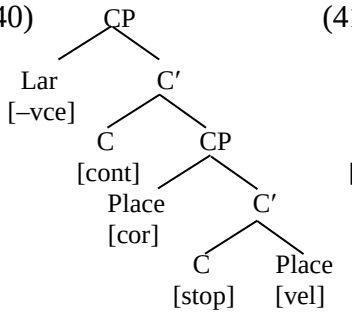
[49] If phonation types (breathy voice, creaky voice) represent laryngeal properties and can be phonologically contrastive in Nuclei, then two possibilities present themselves. In syllables without Onsets, phonation can be mapped into Spec*v*P. In syllables that do have an Onset (and whose Spec*v*P is hence taken), phonation can be represented as a secondary articulation, with the aid of adjunction.

Spec*v*P position is occupied by the Onset, cannot also have a laryngeal specifier, hence cannot be contrastively specified for laryngeal information.

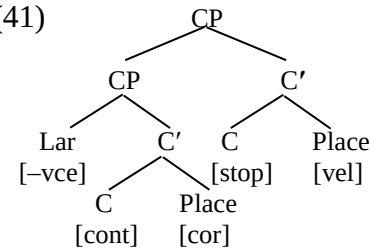
5.2 Onset clusters

In the syntax of phonology, clusters of segments are represented with the aid of specification and complementation, with the choice between the two being contextually determined. In observance of sonority sequencing, an /sk/ cluster in Coda position has a structure representing /s/ as the head and that of /k/ in its complement (as in (40)), whereas /sk/ in onset position has the structure for /s/ in the specifier position of the structure for /k/ (as shown in (41)).

(40)



(41)



Note that in (41) the /s/ is represented as the specifier of the plosive. Because laryngeal information for /k/ is ordinarily represented in the plosive's specifier position, in /sk/ Onset clusters it is impossible to specify the fricative and the plosive separately for laryngeal information: the laryngeal specification for the /sk/ onset cluster is housed in the specifier position of the fricative. This derives the fact that in English /sk/-onsets, the /k/, even when immediately preceding the nucleus of a stressed syllable, is not aspirated: aspiration is not a feature for which fricatives are specifiable in English; since the fricative is the host of the laryngeal features for the entire /sk/ onset cluster, it follows that /k/ cannot be aspirated in this environment.

In the Coda cluster structure in (40), the plosive is again not specified for laryngeal information of its own, this time because the specifier position of the plosive is occupied by the place specification for the fricative. Even though it is no longer a sister of the manner-head of the fricative, the place specification [cor] is still uniquely associated with the fricative: only the fricative manner-head (C[cont]) c-commands this place specification, so only this manner-head can establish an Agree relation with this place feature. Place specification in phonological structures obtains under closest c-command, not under sisterhood. A close syntactic parallel here is with accusative Case assignment, which often seems to take place under sisterhood, but the only generalization that fits the entire bill is one that says that accusative Case is assigned under Agree (i.e., closest c-command). Thus, compare *I considered this proposal*, in which *this proposal* is the verb's sister, with *I considered this proposal interesting*, where the same noun phrase is now the verb's niece (i.e., a daughter of verb's sister), on the plausible assumption that *consider* in the latter example takes a small clause [*this proposal interesting*] as its complement. In the same way in which *this proposal* is 'shifted downward' into a niecehood relation with the verb under the addition of the secondary predicate *interesting*, so also the place specification for the fricative /s/ (which is 'ordinarily' its complement) is 'shifted downward' into a niecehood relation with the fricative's manner-head (C[cont]). The 'integrity of the seg-

ment' can thus be broken, under the influence of the placement of a full X-bar structure in the complement of a head.

In both (40) and (41), there is room for but a single laryngeal specification, harbored by the specifier of the fricative in both cases. The stop does not have space for a laryngeal specification of its own: its specifier position is occupied, in (40) by the place specification of the fricative, and in (41) by the entire structure of the fricative. The fact that the stop cannot itself be specified for laryngeal properties accounts directly for voicing assimilation in clusters of the fricative+stop. A clear connection presents itself here with the work of Kehrein & Golston (2004), and also Golston & van der Hulst (1999) and van der Hulst (in prep.), where it is argued that syllabic units (Onsets, Nuclei and Codas) can have only one laryngeal and place specification.

So far in this discussion of consonant clustering we have confined ourselves to clusters with an initial fricative and a following stop. Such clusters obey the sonority sequencing principle in Coda position (which is what gives rise to the head–complement structure in (40)) but apparently violate it in Onset position. A sonority scale violation is averted, however, by placing the fricative in the specifier position of the plosive in /sk/ onset clusters, as in (41). With this in mind, let us see what the system should say about /ks/ clusters. These obey the sonority scale in onset position but apparently violate it in coda position. Structurally this means that a /ks/ cluster serving as a syllable onset will have the more sonorous element (i.e., /s/) as the complement of the less sonorous element (/k/), whereas a /ks/ cluster in coda position will have the /k/ as the specifier of /s/.

Entirely parallel remarks apply to stop+liquid clusters. So, in an English /kl/ cluster in Onset position, /k/ takes the liquid as its complement. This entails that the laryngeal specification for the cluster is in the specifier position of the structure for /k/. The liquid hosts the place information for the plosive in its specifier, and hence cannot itself be specified for laryngeal properties. The /kl/ Onset cluster has just a single laryngeal specification – the one in the specifier position of /k/, which is the element for [–voice]. This laryngeal specification scopes over the entire cluster. This derives the fact that in stop+liquid onset clusters in English, the liquid is devoiced. For Dutch, which has no aspiration of voiceless plosives in onsets of stressed syllables, the /l/ in /kl/ onset clusters will be voiced by *default*; the voicing of /l/ in this context is not explicitly represented in the structure. More generally, the prediction that this analysis of stop+C sequences in Onset position makes is that the second element should never be *contrastively* specifiable for laryngeal properties, which seems correct: only liquids, nasals and voice-assimilating fricatives occur in second position in such Onset clusters.

5.3 Codas and the place properties of the Nucleus

The complement position of the manner-head is the locus for the specification of place of articulation. The place feature does not necessarily have to *be* the complement of the manner-head; but it does have to be in a 'closest c-command' relation with the manner-head (recall the discussion of (40), above). In syllables whose complement-of-V position is occupied by a (non-sonorant) Coda, this leads to the prediction that the distinctive place-of-articulation properties of the syllable Nucleus will be 'shifted downward' into the specifier position of the Coda consonant in the complement-of-V position. In light of the fact that this specifier position is 'ordinarily' the locus of the laryngeal specification of this consonant, this leads to the

expectation that a Coda consonant in the complement-of-V position which has to harbor the place specification for the Nucleus cannot be contrastively specified for laryngeal properties.

This delivers a simple perspective on 'final devoicing' in languages such as Dutch or German. When a non-sonorant consonant serves as the Coda of a closed syllable, this consonant is necessarily deprived of voicing, and surfaces voiceless. This follows since, sitting in the complement-of-V position, this consonant must harbor the place feature of the Nucleus, and can itself only have the unmarked value for voicing, which in Dutch and German is [–vce].

For languages (such as English) which do not have final devoicing, the most straightforward interpretation of the facts, from the perspective of our proposal, would be that their non-sonorant Codas are only *apparent* Codas: structurally, they are mapped into the Onset position of a following syllable (with a silent Nucleus).

5.4 Adjunction: Nasality, tone, secondary articulation

Beyond the head, specifier, and complement positions, additional distinctions can be made with the aid of another mechanism familiar from phrase-structure syntax: *adjunction*. Adjunction is a useful tool for making the oral/nasal distinction. When nasality is strictly confined to an individual segment (for instance, only to the vowel nucleus), adjunction takes place directly at the level of the head. But the nasality marker can also be adjoined higher up the tree. By exploiting the level of adjunction, we can account for the 'reach' of the nasal property (thus, Golston & van der Hulst (1999:156) point out that nasality can associate to the entire syllable).

For tone, an approach in terms of adjunction also suggests itself, especially for 'spreading' tonal autosegments: adjuncts can have scope over a large portion of the structure; the higher they are adjoined, the wider their scope. Secondary articulations are naturally expressed in the structure with the help of adjunction as well. We will see this at work in the following subsection, where we revisit the place-of-articulation restriction on Coda clusters consisting of a sonorant and a stop, brought up previously in section 3.5, to fill in the details.

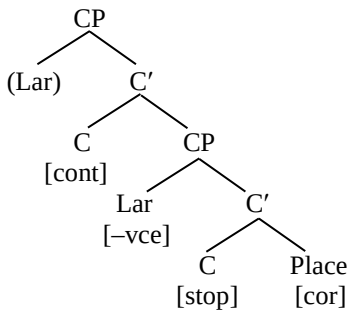
5.5 Coda clusters and place of articulation

We have come across a few situations in which both a sonorant and a stop had to be accommodated in the complement-of-V position (as in English *wield*), and we have seen that such situations impose severe place-of-articulation restrictions on the Coda cluster. To understand these properly, we need to consider carefully what the resulting consonant cluster looks like – and for this, an understanding of the internal structure of consonants, along the lines of (4b), is highly revealing.

In sonorant+stop sequences which are mapped into the complement-of-V position, the stop is in the complement position of the sonorant. This, in conjunction with the fact that the complement position of a consonant is where its place of articulation is specified, entails that the sonorant of a Coda cluster cannot be specified *by itself* for place. The structure in (42), for a cluster such as /lt/, makes this immediately clear:[50]

[50] We placed 'Lar' in the highest specifier position in parentheses because laryngeal information is non-contrastive in the case of liquids, hence arguably phonologically represented only when marked.

(42)



In this structure, the liquid+stop sequence has but a single specification for place: that of the stop, with which the sonorant agrees. This specification must be coronal: otherwise, the liquid, whose C-head c-commands the place specification in the complement of the stop, would be unpronounceable. This is how we derive the fact that the liquid+stop sequence following a long vowel or diphthong in English must be coronal (*wield* versus **wielk*).

For nasal+stop sequences in Coda position, Botma *et al.* (2008: sect. 7) find a picture similar to the one documented for liquid+stop sequences in this position. Again, when the preceding Nucleus is a long vowel or a diphthong, nasal+stop Coda clusters can only be coronal (at least in monomorphemic words[51]): the only allowable such sequences are coronal (*fiend*, *find* and *wont* 'inclined', but not **liemp*, contrasting with *limp*, with a short vowel). English nasals ordinarily have a variety of different places of articulation. Why can't these all be shared equally by the two C's in the cluster in (42)? To see this, we need to understand place specification in a bit more detail.

For simplicity of initial illustration, we had used the features [cont], [vce], [stop] and [cor] in our structure in (42). But we believe that the entire *SPE*-based system of distinctive features can and should be replaced with a system in which, throughout the entire range of contrasts, only two unary features or 'elements' are used: |V| for 'open', and |C| for 'closed'. The interpretation of these features depends on the location of these features in the hierarchical structure. For consonants, in the manner of articulation domain (the head of the structure in (4b)) |V| 'open' translates as 'continuant' and |C| as 'stop'; and in the place system (the complement position), |V| means 'low and back', and |C| as 'high and front'. To represent a coronal consonant in this system, a |C| in complement position, marking 'high and front', is sufficient all by itself. But a velar consonant cannot just be marked for place by |V|, which means 'low and back': the 'back' component is certainly appropriate, but to ensure a velar output, the feature |C| must be added as a modifier of |V|, to raise the place articulation up to the velum. For labial consonants, too, a representation involving adjunction is needed: labial(ization) is a secondary articulatory gesture, marked once again as |C| (closing of the lips) in an adjoined

---

[51] We added this parenthesis because, as is well known, long vowels and diphthongs can perfectly well be followed by a sequence of a non-coronal nasal and /d/ when this /d/ represents the past-tense or past-participial morpheme: *seemed*, *claimed* and *rhymed* are cases in point. For such sequences, morpheme-to-syllable homomorphism may lead to a structure in which there are two syllables present, the second one representing the inflectional morpheme. The nasal is mapped into the Onset position of the second syllable, and the stop /d/ forms its Coda, occupying the complement-of-V position by itself. (The Nucleus of the second syllable is silent in the examples quoted above; but under the right circumstances, which we will not attempt to characterize, this Nucleus is spelled out as schwa or /ɪ/, as in *learnéd society*.)

We note for full disclosure that Dutch does allow tautomorphemic sequences of a non-coronal nasal and a coronal stop following a long tense vowel: /fre:mt/ *vreemd* 'strange'.

position, this time to the element |C|. The details regarding the representation of velar and labial consonants need not concern us. What is important for our purposes is that while plain coronal consonants such as /t/ have their place of articulation specified by a bare element |C| in complement position, velar /k/ and labial /p/ require the presence of a *modifier* in an *adjunction* position to the basic place element.

Let us now return to the structure in (42). Sonorant+stop sequences in the complement-of-V position force the sonorant and the stop into sharing a single place specification in the complement of the stop, under Agree (or c-command). For coronals, this is straightforward: the C-head representing the liquid can engage in a 'long-distance agreement' relation with the place feature |C| in the stop's complement. Now consider velar and labial nasal+stop sequences. Here, the Agree relation between the nasal and the place specification of the stop in (42) fails to fully specify the nasal for the same place of articulation as the stop, which is complex, involving an adjunction structure. This causes the result to crash. In the case of a *coronal* nasal+stop sequence, by contrast, Agree specifies the nasal in just the right way: both nasal and stop are specified as 'plain' |C|, interpreted as 'high and front' (i.e., [coronal]). This explains why tautomorphemic nasal+stop sequences in Coda position following a long vowel or diphthong, where these sequences must be mapped into the complement-of-V position, can only be coronal.

6 Reflections on why recursion is more pervasive in syntax than in phonology

The central thesis of this chapter is that phonology and syntax have recourse to the same computational system, i.e. that both modules are maximally analogous. This thesis goes beyond the claim that both phonology and syntax build hierarchical structures. This claim is commonly made (though not supported by all linguists) with the proviso that the nature of the hierarchical organization is fundamentally different with phonology adhering to 'strict layering', while syntax displays recursive structure. Accepting that recursion is available to phonology does not entail that phonology will display the same amount of recursive structure as morphosyntax. The kinds of structures that are employed in both modules do not exist in a vacuum, but rather are formed to accommodate the substances that these structures are grounded in.

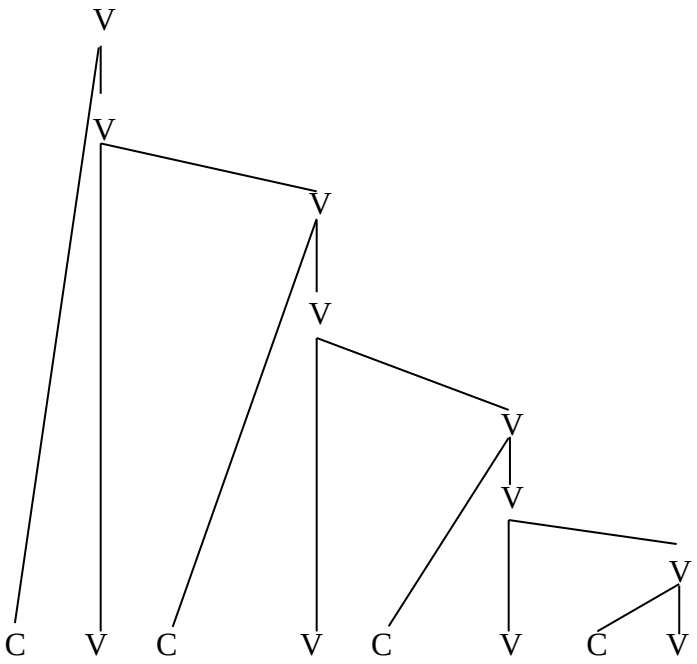
We have already pointed out that syntax displays more syntactic structure than phonology due to the lack of a parallel to morphosyntactic functional categories in the latter. However, there is an additional reason for why recursion in phonology is less pervasive. If we accept the fact that semantic, conceptual structure (Anderson would say 'conceptual *substance*') is inherently recursive, we expect morphosyntax be isomorphic to this semantic, conceptual structure as much as is possible. Certain factors that cause syntactic displacements of various kinds entail a lack of isomorphism, creating a mismatch between morphosyntactic structure and semantic-conceptual structure, which testifies to the relative autonomy of the two modules. Phonological structure accommodates phonetic-perceptual substance, which arguably is not inherently recursive. Rather, as the result of motoric actions, it is essentially sequential.

This may lead to a view that phonology is 'flat' (see Scheer 2013), perhaps only displaying recursion when expressions are morphosyntactically structured. But recursion in phonology is limited even in this case because there is a 'flattening force' that causes disrhythmic structures that contain lapses (sequences of weak units, 'SWWW…') to flatten by breaking up in smaller rhythmic units (i.e. SW SW), as shown in Giegerich (1985). This in

itself shows that phonological structure is not entirely flat. After all, if there is rhythmic structure this means that the units (syllables, words, etc.) display a structure in which certain units are 'subordinated' to others. Standard metrical phonology has chosen to formally represent this 'subordination' by grouping units into binary, headed constituents. The crucial point of van der Hulst's (2010b) proposal was that subordination can also be encoded in terms of embedding, which then establishes a perfect formal parallel with recursion in syntax.

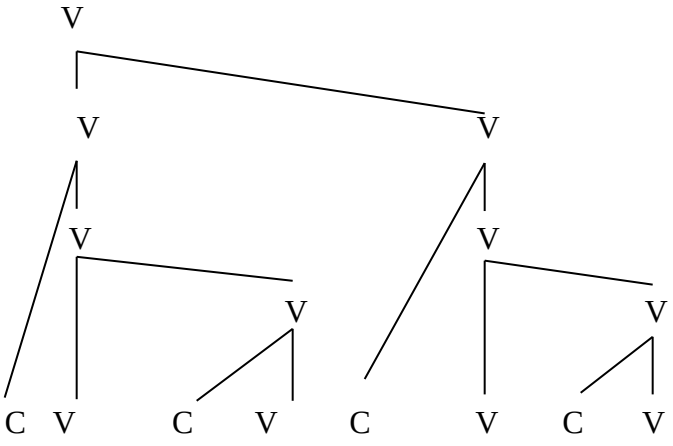
But the same flattening forces that limit phonological recursion in morphosyntactically structure expressions also prevents level-3 embedding in monomorphemic units. A sequence of four syllables is therefore not structured as a quaternary 'foot'.

(43)



Although formally perfectly correct, (43) creates a dysrhythmic sequence SWWW that does not match the rhythmic structure of a quadrisyllabic sequence. Indeed, a string of four CV units is likely to display an alternating rhythmic structure (SWSW), which suggests the presence in the structure of two consecutive units, each with level embedding:

(44)



(44) is 'flatter' than (43) and this, we suggest (merely making explicit what most phonologists would take for granted), is a consequence of the rhythmic nature of the 'phonetic substance' that phonotactic structure represents. Beyond the 'magic number' 3, unbounded recursion gives in to rhythm.

7 Conclusion

In this chapter, we have explored a 'radical' approach to the structural analogy assumption. Rather than making suggestions for parallelism based on a 'naïve' version of syntactic theory, we have investigated in detail potential uses in both phonology and syntax of mechanisms that are standardly thought of as being exclusively syntactic, such as recursion, X-bar structure, and, more specifically, the 'light *v*' structure of multi-argument constructs. We have seen that assigning subordination structures to phonology – not just at the level of the foot but also within the syllable and even in the representation of segments – opens up explanatory perspectives on many a persistent question.

One question that this leaves us with is why, if recursion in phonology is curtailed to a depth of embedding that does not go beyond a structure that is dactylic, phonology could not be limited to adjunction (rather than subordination). If the computational system that is available to phonology and syntax makes both subordination and adjunction available to accommodate apparently recursive effects, why would phonology not limit itself to adjunction? This question boils down to asking whether subordination or adjunction is the default option. If recursion is, as Chomsky now argues, 'The Basic Property' of language, we would be inclined to take subordination as the default mechanism. The usefulness of subordination in phonology reveals itself at many different levels, as we have shown. None of the more microscopic predictions (including but not restricted to those made in connection with 'segment integrity') would be made by a model confining itself to adjunction as the combinatory mechanism in phonology. Recursion in phonology is real, and its results are revealing.

References

Anderson, John. 1987. The tradition of structural analogy. Language topics: essays in honour of Michael Halliday, ed. by R. Steele & T. Threadgold, 33-43. Amsterdam: John Benjamins.

Anderson, John. 2011 The substance of language, Volume III: Phonology-syntax analogies Oxford: Oxford University Press.

Baker, Mark. 1988. *Incorporation. A theory of grammatical function changing*. Chicago: University of Chicago Press.

Belletti, Adriana. 2004. Aspects of the low IP area. In Luigi Rizzi (ed.), *The Structure of CP and IP. The Cartography of Syntactic Structures*, Vol.2. New York: Oxford University Press, 16-51.

Bickerton, Derek. 2000. Calls Aren't Words, Syllables Aren't Syntax. Review of Carstairs-McCarthy on Language-Origins. Psycoloquy: 11(114).

Bošković, Željko. 1997. *The syntax of non-finite complementation*. Cambridge, MA: MIT Press.
Botma, Bert, Colin Ewen & Erik Jan van der Torre. 2008. The syllabic affiliation of postvocalic liquids: An onset-specifier approach. *Lingua* 118: 1250-70.
Bromberger, Sylvain & Morris Halle. 1989. Why phonology is different. *Linguistic Inquiry* 20 (1): 51-70.
Carr, P. 2006. Universal grammar and syntax/phonology parallelisms. Lingua 116: xxx-xxxx.
Carstairs-McCarthy, Andrew. 1999. *The origins of complex language. An inquiry into the evolutionary beginnings of sentences, syllables and truth*. Oxford: Oxford University Press.
Chomsky, Noam. 1994. Bare phrase structure. MIT Occasional Papers in Linguistics. Cambridge, MA: MIT Department of Linguistics and Philosophy, MITWPL.
Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.
Dikken, Marcel den. 2006. *Relators and linkers*. Cambridge, MA: MIT Press.
Dikken, Marcel den & Rint Sybesma. 1998. *Take*-serials light up the middle. Paper presented at GLOW Tilburg; ms., CUNY Graduate Center & University of Leiden.
Fudge, Eric. 1987. Branching Structure within the Syllable. Journal of Linguistics, Vol. 23, No. 2 (Sep., 1987), pp. 359-377.
Garcia-Bellido, Paloma. 2005. The morphosyntax and syntax of Phonology: The svarabhakti construction in Spanish, *Estudios de Lingüística del Español*, Vol. 22.
Giegerich, H.J. 1985. *Metrical Phonology and Phonological Structure*. Cambridge University Press, Cambridge.
Golston, C. & H. van der Hulst. 1999. Stricture is structure. In: B. Hermans & M. van Oostendorp (eds.). *The derivational residue in phonological Optimality Theory*. Amsterdam: John Benjamins, 153-174.
Gordon, M. 1998. The phonetics and phonology of non-modal vowels: a crosslinguistic perspective', *Berkeley Linguistic Society* 24, 93-105.
Gussenhoven, Carlos. 2008. Vowel Duration, Syllable Quantity, and Stress in Dutch. In: Kristin Hanson and Sharon Inkelas. Eds. The Nature of the Word. The MIT Press, 181-198.
Hale, Kenneth & S. Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & S. Jay Keyser, eds., *The View from Building 20*, 53–104. Cambridge, MA: MIT Press.
Hayes, B. (1995). *Metrical Stress Theory: Principles and case studies*. Chicago: University of Chicago Press.
Hulst, Harry van der. 1984. *Syllable structure and stress in Dutch*. Dordrecht, Foris.
Hulst, Harry van der. 1985. Ambisyllabicity in Dutch, In: H. Bennis & F. Beukema (eds.). *Linguistics in the Netherlands 1985*. Dordrecht: Foris, 57-67.
Hulst Harry van der. 1996. Separating primary accent and secondary accent. In: R. Goedemans, H. van der Hulst & E. Visch (eds.). *Stress patterns of the world. Part I.* HIL Publications 2. The Hague: Holland Academic Graphics, 1-26
Hulst, Harry van der. 1997. Primary accent is non-metrical. *Rivista di Linguistica 9/1*, 99-127 Hulst, Harry van der. 2005. The molecular structure of phonological segments. In *Headhood, elements, specification and contrastivity*. P. Carr, J. Durand and C. Ewen (eds.). 193-234. Amsterdam: John Benjamins Publishing Company.

Hulst, Harry van der. 2008. The Dutch diminutive. In: C. Ewen, H. van der Hulst & N. Kula (eds.). *Trends in prosodic phonology*. *Lingua* volume 118, issue 9, 1288-1306.
Hulst, Harry van der. 2010a. Re Recursion. In: Harry van der Hulst (ed.). Recursion and Human Language. Berlin: Mouton de Gruyter, xv-liii.
Hulst, Harry van der. 2010b. A note on recursion in phonology. In: Harry van der Hulst (ed.). Recursion and Human Language. Berlin: Mouton de Gruyter, 301-342.
Hulst, Harry van der. in prep. Principles of Radical CV Phonology.
Hulst, Harry van der. 2012. Deconstructing stress. *Lingua* 122, 1494-1521.
Hulst, Harry van der and Nancy Ritter. 1998. Kammu minor syllables in Head-driven Phonology. In: E. Cyran (ed.). *Structure and interpretation. Studies in phonology*. Lublin: Folium, 163-182
Hunyadi, Laszlo. 2010. Cognitive grouping and recursion in prosody. In: Harry van der Hulst (ed.) *Recursion and human language*. Berlin: Mouton de Gruyter, xxx-xxx
Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology Yearbook* 7: 193-231.
Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Kehrein, W. and C. Golston. (2004). A prosodic theory of Phonation contrasts. *Phonology* 21, 1–33.
Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring, eds., *Phrase Structure and the Lexicon*, 109–137. Dordrecht: Kluwer.
Kurylowicz, J. 1948. Contribution a la theorie de la syllabe. BPTJ 8. 80-114.
Ladd, D. Robert. 1996. *Intonational Phonology.* Cambridge: Cambridge University Press. [2nd edition 2008]
Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19: 335–391.
Levin, J. (1985). *A metrical theory of syllabicity*. Doctoral dissertation, MIT.
Martínez-Paricio, V. 2013. An exploration of minimal and maximal metrical feet. PhD dissertation, CASTL, University of Tromsø.
Mutlu, Filiz. 2017. Valence and saturation in phonology. Master Thesis Boğaziçi University.
Neeleman, Ad, and Hans van der Koot. 2006. On syntactic and phonological representations. *Lingua* 116/10, 1524-1552.
Pike, K.L. & E.V. Pike. 1947. Immediate constituents of Mazateco syllables. *IJAL* I3.
Pinker, Steven and Ray Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95: 201–236.
Polgárdi, Krisztina. 2008. The representation of lax vowels in Dutch: A loose CV approach. *Lingua* 118: 1375-92.
Rowicka, Grazyna. 1999. On trochaic Proper Government. In: Rennison, J., Kühnhammer, K. (Eds.), *Phonologica 1996: Syllables!?* Holland Academic Graphics, The Hague, pp. 273–288.
Scheer, Tobias. 2013. Why phonology is flat: The role of concatenation and linearity. Paper presented at the 11th Rencontres du Réseau Phonologique Français Nantes 1-3 juillet 2013
Smith, N.S.H. 1999. A preliminary account of some aspects of Leurbost Gaelic Syllable structure. In H. van der Hulst & N. Ritter (eds), *The syllable: views and facts.* Berlin: Mouton de Gruyter, 557-630.
Smith, N.S.H. 2003. Evidence for recursive syllable stuctures in Aluku and Sranan. In: N.S.h. Smith and D Adone - Recent development in creole studies, 2003.

Szigetvári, Péter. 2016. No diphthong, no problem. In: Jolanta Szpyra-Kozłowska and Eugeniusz Cyran (eds.) *Phonology, its faces and interfaces.* Frankfurt am Main: Peter Lang. 123–141.

Tallerman, Maggie. 2006. Challenging the syllabic model of 'syntax-as-it-is'. *Lingua* 116: 689–709

Trommelen, M. 1983. *The syllable in Dutch: With special reference to diminutive formation*. Dordrecht : Foris.

Ulfsbjorninn, S. 2015. A Field Theory of Stress. Ph.D. Dissertation. SOAS – University of London.

Vijver, Ruben van de. 1998. *The Iambic Issue: Iambs as a Result of Constraint Interaction*. Amsterdam, The Netherlands: Vrije Universiteit dissertation.

Völtz, Michael. 1999. The syntax of syllables: why syllables are not different. In: John Rennison and Klaus Kühnhammer (eds.), *Phonologica 1996*. The Hague: Holland Academic Graphics, 315-321.

Wagner, Michael. 2005. Prosody and recursion. Ph.D. diss., Massachusetts Institute of Technology, Cambridge, MA.

Weijer, Jeroen van de and Jisheng Zhang. 2008. An X-bar approach to the syllable structure of Mandarin. *Lingua* 118: 1416–1428.