The extended prosodic word

Junko Ito and Armin Mester

1. Introduction

In an extended prosodic word, a word-sized unit $\alpha$ has combined with additional material $\beta$ to form, by adjunction, a larger recursive structure $[\omega[a][\beta]]$ or $[\alpha[\beta][\omega]]$. Such structures have often been posited, following Inkelas (1989) and others, when $\beta$ is a loosely bound derivational or inflectional suffix, replicating (more or less precisely, due to the pressures of syllabification constraints) morphological structure in prosody. This is a way of capturing the closure effects associated with the word cycle studied by Borowsky (1993) and others, resulting in $\text{sing}^{\#er}$ without $[g]$ (vs. $\text{finger}$ with $[g]$), trisyllabic $\text{kind}^{[[\#]]\text{ing}}$ (vs. disyllabic $\text{kindling}^{[\#]}$), etc. Even though sometimes seen as a virtue, the duplication here is worrisome. Paradigm constraints and other Output-Output (OO) conditions, an important and widely accepted part of current Optimality Theory (OT), cover much the same ground, and the idea of writing closure effects into the phonological representation itself raises the suspicion that the structures posited are little more than a legacy from the days of the heavily representation-focused approach to phonology of the 1970’s and 80’s, a way of encoding effects that are better accounted for elsewhere in the grammar, by non-representational means.

We are in basic agreement with this view, which removes much of the motivation for recursive prosodic tree structure that simply mirrors morphological form. What, then, is the status of extended prosodic words, and of adjunction in prosody? We argue that they still have a significant role to play, not as replicas of morphological structure within phonology, but rather at the post-lexical level, as one important option for the parsing of functional items, such as prepositions and determiners. This claim, which takes up earlier proposals by Booij (1996), Vigário (1999), and others, is in conflict with the direct phrase attachment of functional material favored by Selkirk (1996), whose influential work was followed by many researchers. The two opposing views are juxtaposed in (1).

(1) a. phrase-attached: 
   $[\_\text{the [dinosaurs]}]$ 
   b. word-adjointed: 
   $[\omega\text{the [dinosaurs]}]$
The difference might appear small, but it has far-reaching consequences that this paper sets out to explore. Besides the status of the function word itself, an important point of divergence between the two structures lies in the prosodic status of the lexical word dinosaurs: It is a full and independent prosodic word in [the dinosaurs], but not in [the dinosaurs], where it is only a segment of a larger prosodic word. In terms of the approach developed in Ito and Mester (to appear), dinosaurs is a maximal projection of ω in (1a), but not in (1b). This makes a difference to the extent that there are processes that are specific to maximal prosodic words. Insofar as the evidence to be reviewed here is concerned, we will see that the ω-adjoined structure offers the desired structures for English and German.

2. Basic properties of prosodic representation

A significant advance in the study of phonological form since the mid-1980's has been the discovery that speech has its own kind of constituent structure, related to, but in many ways non-isomorphic with, the syntactic structure of utterances. In this general area, a specific line of research has emerged as the main-stream approach, based on precursors including Halliday (1960) and Pike (1967), and developed into a full-scale model in the work of Selkirk (1981), Nespor and Vogel (1986), and others. It conceives of phonological constituent structure as a hierarchy of a small number of well-defined prosodic constituents along the lines of (2), the so-called prosodic hierarchy.

(2) Prosodic Hierarchy: utterance (u)
    |
    intonational phrase (i)
    |
    phonological phrase (ϕ)
    |
    phonological word (ω)
    |
    foot (f)
    |
    syllable (a)

The hierarchy sometimes contains another level subordinate to the syllable, the so-called “prosodic skeleton”, consisting of moraic (µ) or segmental (X, CV) units and expressing quantity and weight distinctions. Following Ito and Mester (2003 [1992]) and McCarthy and Prince (1993a), among others, we assume that skeletal units are not part of the prosodic hierarchy, and that quantity and weight are properties of syllables and segments, not prosodic constituents (see Ito and Mester 2003 for discussion.)

While there are open questions, both of detail and of principle, regarding the exact nature and number of the categories above the prosodic word — indeed, regarding their very existence as separate categories, each with its own label and associated set of defining properties, as opposed to a simple string of unlabeled nodes in a metrical tree structure (Wagner 2005) — it is here convenient to take something like (2) as a broadly shared starting point.

Prosodic representations are instances of the tree structures familiar from syntax and other parts of linguistics. Formally speaking, we can conceive of a prosodic representation as a directed graph consisting of a finite set of nodes, where each node is an ordered pair whose first member is a sequence of natural numbers <α>, its address, and whose second member is a label L. An example is (3), given both in its algebraic form (a) and in a representation where precedence and dominance are graphically expressed (b). The latter is usually simplified to (c).

(3)    a. {<(1>, ω), (<1,1>, f), (<1,2>, f), (<1,1,1>, σ), (<1,1,2>, σ), (<1,2,1>, σ), (<1,2,2>, σ)}

       b. (1>, ω)
         /     /
       (1>, f) (1,2>, f)
         /     /
       (<1,1>, σ) (<1,2>, σ)
         /     /
      f     f
         /   /
      σ   σ   σ

The (unique) root of the tree has the address <1>; if a given node has the address <α>, its immediate daughters have addresses <α,1>, ..., <α,n>. Node A dominates node B if A's address is <α> and B's address is <α,β>; if, in addition, β = i, for some number i, A dominates B immediately. Given a pair of nodes, A with address <α,i>, and B with address <α,j> such that j > i, A and every daughter of A precedes B, and every daughter of B. This means that prosodic trees are non-tangling in the sense of Wall (1972: 148).
Prosodic trees have a number of basic properties, as listed in (4).

(4) Basic properties of (well-formed) prosodic trees (see Wall 1972: 144–152 for the formal background)

a. **Rootedness**: There is exactly one node that dominates every other node.

b. **Linear order**: The nodes immediately dominated by a node are linearly ordered from left to right.

c. **No tangling**: For any nodes x and y, if x precedes y, then all nodes dominated by x precede all nodes dominated by y. This excludes both line crossing and improper bracketing (since nodes do not precede themselves).

d. **Labeling**: Each node bears a label, an element of the ordered set \( \text{PH} = \{ U > I > ? > w > f > a > p \} \), the prosodic hierarchy, whose elements stand in a relation of containment, as indicated.

e. **Containment**: Each immediate dominance relation respects the containment structure of the prosodic hierarchy, in the sense that lower-ranked elements do not immediately dominate higher-ranked elements.

f. **Headedness**: Every (non-terminal) prosodic category dominates a head, a prosodic category at the next lower level in the prosodic hierarchy.

In an optimality-theoretic conception of phonology, here assumed throughout, it is convenient to assign the conditions in (4) to GEN, given their basic nature and the fact that there is little reason to suspect any of them to ever be violated. They are no more than a minimum consensus on prosodic representations and admit a wide variety of structures, many of them never encountered. What else needs to be assumed in order to more precisely characterize the form of prosodic representations? Building on previous work, we assume four kinds of requirements, which we will take up one-by-one in what follows.

(5) Constraints on prosodic representations

a. **Parsing constraints**

b. **Interface conditions**

c. **Size and shape requirements**

d. **Tree form restrictions**

2.1. **PARSE-INTO-X**

Prosodic parsing is the assignment of constituent structure at each of the levels of the prosodic hierarchy. In OT the basic parsing imperative is implemented as a set of specific constraints, often crucially dominated, requiring parsing at each level. This is the strategy taken by Prince and Smolensky (2004 [1993]: 64), who refer to “PARSE-\( \sigma \), which must be distinguished from the other members of the PARSE-element family”, such as PARSE-segment (Prince and Smolensky 2004, note 40). In another context, they note “the effects of a principle of exhaustive metrical analysis, familiar from the earliest work in the area [. . .]. This principle is part of the parsing theory, and we will call it PARSE-\( \sigma \), omitting from the name the information that \( \sigma \) is parsed into F” (Prince and Smolensky 2004: 62). Overall, the family of PARSE-element constraints plays a fundamental and ubiquitous role in their work. Extending this approach, with some modifications, we conceive of prosodic parsing as driven by a family of OT constraints PARSE-INTO-X, where \( X \) is an open parameter standing for the levels of the prosodic hierarchy (2).

(6) **PARSE-INTO-X**:

\{PARSE-INTO-0, PARSE-INTO-f, PARSE-INTO-\( \omega \), PARSE-INTO-\( \varphi \), PARSE-INTO-I, PARSE-INTO-U\}

"Every element of the terminal string is parsed at the X-level"

The elements of the terminal string are the phonological segments. PARSE-INTO-X only requires that segmental strings belong to \( X \), and is crucially not "PARSE-INTO-WELLFORMED-X". For example, PARSE-INTO-f and PARSE-INTO-\( \sigma \) are themselves not concerned with the wellformedness of feet and syllables. These units have their own intrinsic defining properties stated as separate constraints (for feet, binarity and rhythmic requirements; for syllables, sonority requirements, coda restrictions, complexity, etc.).
Full parsing can be seen in an example like (7). We often use equivalent bracketed representations, as in (8), where periods mark syllable boundaries, parentheses show foot boundaries, and brackets with subscripts indicate higher-level constituents.

(7) \[ \left[ \left[ \left[ \left[ \left[ \text{'di} \text{nosaurs} \right] \right] \right] \right] \right] \right] \left[ \left[ \left[ \left[ \left[ \left[ \text{roamed} \right] \right] \right] \right] \right] \right] \]

(8) \[ \left[ \left[ \left[ \left[ \left[ \left[ \text{Di no saurs} \right] \right] \right] \right] \right] \right] \left[ \left[ \left[ \left[ \left[ \left[ \text{Ani zo na} \right] \right] \right] \right] \right] \right] \]

The constraints at every level is enforced by the family of Parse-into-X constraints in (6). To see how these constraints work, consider various parsings of the utterance dinosaurs! given in (9): The totally unparsed (a), syllabified (b), syllabified and footed (c), etc. The fully parsed (g) shows all the levels of the prosodic hierarchy. Given the simple formulation of the parsing constraints in (6), the target is in each case the terminal string of segments, and it is at this level that violations are reckoned, as shown in (9).

(9) \[
\begin{array}{|c|c|c|c|c|c|c|}
\hline
& /\text{di} \text{nosaurs}/ & \text{Parse-into-o} & \text{Parse-into-f} & \text{Parse-into-\( \omega \)} & \text{Parse-into-\( \varphi \)} & \text{Parse-into-\( \iota \)} \\
\hline
a. & /\text{di} \text{nosaurs}/ & \text{damas} & \text{damas} & \text{damas} & \text{damas} & \text{damas} \\
\hline
b. & /\text{di} \text{nosaurs}/ & \text{damas} & \text{damas} & \text{damas} & \text{damas} & \text{damas} \\
\hline
c. & /\text{di} \text{nosaurs}/ & \text{damas} & \text{damas} & \text{damas} & \text{damas} & \text{damas} \\
\hline
d. & /\text{di} \text{nosaurs}/ & \text{damas} & \text{damas} & \text{damas} & \text{damas} & \text{damas} \\
\hline
e. & /\text{di} \text{nosaurs}/ & \text{damas} & \text{damas} & \text{damas} & \text{damas} & \text{damas} \\
\hline
f. & /\text{di} \text{nosaurs}/ & \text{damas} & \text{damas} & \text{damas} & \text{damas} & \text{damas} \\
\hline
g. & /\text{di} \text{nosaurs}/ & \text{damas} & \text{damas} & \text{damas} & \text{damas} & \text{damas} \\
\hline
\end{array}
\]

The constraints are unranked in this tableau (indicated by vertical broken lines), and the less structure the candidates have at the various levels, the less harmonic they are. There are also output candidates where parsing has skipped a level, either partially or entirely, as illustrated in (10).

(10) Partially parsed candidates

<table>
<thead>
<tr>
<th>/\text{di} \text{nosaurs}/</th>
<th>Parse-into-o</th>
<th>Parse-into-f</th>
<th>Parse-into-( \omega )</th>
<th>Parse-into-( \varphi )</th>
<th>Parse-into-( \iota )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. /\text{di} \text{nosaurs}/</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
</tr>
<tr>
<td>b. /\text{di} \text{nosaurs}/</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
</tr>
<tr>
<td>c. /\text{di} \text{nosaurs}/</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
</tr>
<tr>
<td>d. /\text{di} \text{nosaurs}/</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
</tr>
<tr>
<td>e. /\text{di} \text{nosaurs}/</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
</tr>
<tr>
<td>f. /\text{di} \text{nosaurs}/</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
</tr>
<tr>
<td>g. /\text{di} \text{nosaurs}/</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
<td>damas</td>
</tr>
</tbody>
</table>

Candidate (10b) does not parse its segments into \( \omega \), and (10c) does not parse them into \( f \) (feet). There can also be candidates where parsing into a level is partial, as illustrated in (10d) for \( f \). In this example, none of these partially parsed candidates is of much interest, since the fully parsed (10a) is the clear winner, fulfilling all Parse-into-X constraints. Depending on the segmental make-up, however, there can be situations where the non-fully parsed candidate will be the winner. This is a result of interactions with other constraints, as will be shown next.

2.2. Interactions of Parse-into-X with other constraints

A basic question concerns the fact that prosodic structure is generally maximally articulated. Since Parse-into-X demands only that the terminal string be dominated by \( X \), why build more than the minimally required single \( X \) for any reason? For example, why does the whole sentence not become one gigantic prosodic word? This question does not arise in the same way at the lower levels because there are well-known substantive constraints on feet and syllables that prevent oversized constituents. The whole sentence cannot be one gigantic foot because constraints on foot form allow maximally two syllables per foot (FootBINARITY (FRBIN): “Feet must be binary under syllabic or moraic analysis”, McCarthy and Prince 1993b: 46). Similarly, forming one gigantic syllable is out of the question because constraints on syllable form and sonority sequencing prevent incorporating a consonant into a syllable after a vowel as a coda, and then adding a second vowel to the same syllable. However, at higher levels of structure, such answers are not readily available. Consider a parse like (11) (where syllable and foot structure is assumed to be present).

(11) \[ /\text{Dinosaurs roamed Arizona}/ \]

All Parse-into-X constraints are fulfilled in (11), and English words can clearly contain more than seven syllables, as shown by formaldehydetetramethylamido-
fluorimum, or the notorious floccinaucinifiliplication 'estimation of something as worthless'. So what prevents a parse as in (11)?

The answer lies in constraints on the syntax-phonology interface demanding, generally speaking, that a substring that forms a (certain kind of) syntactic constituent must simultaneously also form a (certain kind of) prosodic constituent.\(^2\) Prince and Smolensky (2004: 45) propose a family of constraints \(Lx \triangleleft Pr\) (MCat) that directly requires a member of the morphological category \(MCat\) to correspond to a prosodic word. A slightly different approach, originated in Chen (1985, 1987) and Selkirk (1986) and later generalized in McCarthy and Prince (1993a), Selkirk (1996), and Truckenbrodt (1995, 1999), sees the imperative of correspondence as a more indirect one: One edge, left or right, of certain syntactic constituents must match an edge of a specific prosodic constituent. Taking up this line of analysis and writing \(Lex\) for "lexical (non-function) word" and \(LexP\) for a corresponding maximal projection, we formulate the relevant interface constraints in (12).

\[(12) \text{LEX-MAPPING} \]
\[\begin{align*}
\text{a. } Lex-TO-\omega(L/R): & \quad \text{Align}(Lex, \text{Left/Right}, \omega, \text{Left/Right}) \\
& \quad \text{"Every lexical word is left/right aligned with a prosodic word."} \\
\text{b. } LexP-TO-\varphi(L/R): & \quad \text{Align}(LexP, \text{Left/Right}, \varphi, \text{Left/Right}) \\
& \quad \text{"Every lexical maximal projection is left/right aligned with a phonological phrase."}
\end{align*}\]

For English, \(Lex-TO-\omega(L)\) and \(LexP-TO-\varphi(R)\) are high-ranking. Tableau (13) shows these constraints in action.

\[(13) \]

<table>
<thead>
<tr>
<th>S</th>
<th>NP</th>
<th>VP</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lex-TO-\omega(L)</td>
<td>LexP-TO-\varphi(R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dinosaurs</td>
<td>roamed Arizona</td>
<td>Dinosaurs</td>
<td>roamed Arizona</td>
</tr>
<tr>
<td>a.</td>
<td>**</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>**</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
<td></td>
</tr>
</tbody>
</table>

\(^2\) We are using "syntactic" here to include both syntactic and morphological structure (i.e., "grammatical", in traditional parlance).

OT constraints are violable, and whether and to what extent they are fulfilled depends on their ranking. The parsing constraints are no exception, and if a constraint on foot form such as \(\text{FTBIN}\) is ranked above \(\text{PARSE-INTO-f}\), non-exhaustive parsing is the result. For example, in (14a) we find a syllable immediately dominated by the word, without an intervening foot (i.e., a violation of \(\text{PARSE-INTO-f}\)). The competing structure (14b), with a foot on [\(ba\)], violates \(\text{FTBIN}\), the higher-ranking constraint on foot form which requires two moras and is violated by the monomoraic ba.

\[
\begin{array}{|l|c|c|}
\hline
\text{banana} & \text{FTBIN} & \text{Parse-into-f} \\
\hline
\text{a. } & \omega & f \\
\text{b. } & \omega & \text{ftBin} \\
\hline
\end{array}
\]

In the analysis of Latin stress in Prince and Smolensky (2004), followed by many others employing exactly this ranking, the tacit assumption is that other \(\text{PARSE-INTO-X}\) constraints are unviolated in the winning candidate. This is then an example of a size constraint ranked above a \(\text{PARSE-INTO-X}\) constraint.

\[
(15) \text{Size Constraint } \gg \text{PARSE-INTO-X} \\
\text{(e.g., FTBIN) } \gg \text{(e.g., Parse-into-f)}
\]

Selkirk (1996) argues that function words in English exhibit \(\text{PARSE-INTO-\omega}\) violations (her \(\text{EXHAUSTIVITY-PHRASE}\)).

We see from (14) that ranking the size constraint \(\text{FTBIN}\) above \(\text{PARSE-INTO-f}\) leads to output structures in which \(\text{PARSE-INTO-f}\) is violated by a single light syllable directly dominated by \(\omega\). There is one situation, however, where such \(\text{PARSE-INTO-f}\) violations lead to ungrammatical results, namely, when the form contains no other material except for the (unstressed) light syllable: A word that just consists of [\(b\)], the first part of [\(ba(\text{nae nemo})\)], is illicit. What is the difference between free-standing [\(\omega ba\)] (16b) and [\(ba\)] in [\(ba(\text{nae nemo})\)] (16a)?
On the empirical side, we know from minimal word effects and similar phenomena in language after language that a single light syllable is often not sufficient to form a prosodic word. The general explanation, originated by Prince (1980), lies in the Headedness requirement (4f), repeated here in (17).

(17) Headedness: Every (non-terminal) prosodic category dominates a head, a prosodic category at the next lower level in the prosodic hierarchy.

Thus Headedness is fulfilled in [be(naenaf)] because \( \omega \) dominates the foot \( (naenaf) \), but it is violated in \([bo]\). The alternative parse, \([(ba)\omega]\), fulfills Headedness but violates FitBin. The combined result of the two constraints is the generalization that a minimal word must be longer than a single light syllable.

The Parse-into-X constraints demand that strings be assigned phonological constituent structure, and Lex-to-\( \omega \) and LexP-to-\( \varphi \) prevent underparsing. What stops strings from being overparsed then? An example would be (18), where the last two syllables are parsed not only as a foot, as in (16a), but in addition as a \( \omega \).

(18) \([\omega bo(\omega naena)]\)

In terms of the parsing constraints, both violate Parse-into-\( \varphi \) in terms of their first syllables, and are tied in this respect. The overparsed (18) contains an extra internal \( \omega \), and the string naena is parsed twice by \( \omega \), once by the internal \( \omega \) and again by the external \( \omega \). With its two \( \omega \)-constituents, one dominating the other, (18) violates a specific tree form constraint, namely, No-Recursion. Different formulations of No-Recursion have been proposed, either as a categorical constraint directly banning self-domination (Selkirk 1996: 190), or as a gradually violable constraint requiring any two prosodic constituents of the same category that are not disjoint in extension to be identical in extension (Truckenbrodt 1999: 240). Continuing the 'string-and-parsing' approach pursued so far, we formulate No-Recursion in (19) as a constraint against multiple parsing of a string by the same category. Informally speaking, (19) militates against one-to-many relations between strings and instances of categories.

(19) No-Recursion: An element is parsed only once into a given category. Assign one violation mark for each additional parse of an element into the same category.

As McCarthy and Prince (1993b) note, recursion of the categories foot and syllable is impossible because "the various principles of foot theory and syllable theory license a very limited set of expansions of foot and syllable," but "there is no theory placing comparable limits on the expansion of the prosodic word". This distinction is also made in Ito and Mester (2007a) in terms of intrinsically versus extrinsically defined prosodic categories. The former are governed by substantive constraints defining their size and shape. Our working hypothesis is that the most basic structural requirements on syllables as well as fundamental rhythmic conditions on feet are universally unviolated (in OT terms, perhaps part of Gen) and preclude recursion. On the other hand, prosodic words and larger prosodic units do not have intrinsically defined shapes but are largely governed by syntax-phonology mapping constraints, which may therefore lead to recursive structures being optimal, depending on the ranking of No-Recursion (19) (see Ladd 1986 for an argument that intonational phrasing can be recursive).

The constraints discussed so far play an important role in determining the structure of complexes consisting of function words and lexical words. Previewing a fuller discussion and motivation in later sections, we note here that the independently motivated constraints give us the following options for such structures.
Making the determiner an independent ω, (20a) violates FrBIN, since a single monomoraic syllable cannot make a foot.4 Fully incorporating the determiner (20b) violates Lex-to-ω, since no prosodic edge marks the left Lex-edge of dinosaur. Adjoining the determiner (20c) to the following ω violates No-RecurSivity; and, attaching the determiner directly to the φ (20d) violates Parse-into-ω.5 Each one of these structures is potentially optimal in some grammar, although some are more likely than others. We will argue that (20c), with its recursive ω-structure, is the correct one for English, contra Selkirk (1996), who has argued for (20d).

2.3. Exhaustivity effects and strict layering

Proposed systems of prosodic hierarchy constraints typically include a third type of constraint whose task is to enforce strict layering. Definitional details aside, “strict layering” describes a situation where full parsing is achieved at every level, so all of the Parse-into-X constraints are fulfilled, as is No-RecurSivity. This was once the standard view of prosodic form. Tentatively suggested, and still remaining nameless, in Selkirk (1981: 381) as something generally true of prosodic representations, the idea acquired a more assertive status in subsequent work, together with its name.

(21)  Strict Layer Hypothesis (Selkirk 1984: 26)
“We have proposed that a category of level i in the hierarchy immediately dominates a (sequence of) categories of level i-1 […] . We will call this the strict layer hypothesis, and will take it as a useful working hypothesis here”.

The same work, however, by adopting a grid-only framework (originated by Prince 1983), abolishes the very core of the prosodic hierarchy – the foot-structure with the prosodic word and the phonological phrase – and replaces prosodic constituent structure by silent grid beats occupying the junctures between words and phrases. With the remaining endpoints of the old hierarchy, the intonational phrase and the syllable, now in immediate domination (see Selkirk 1984: 31), the spirit (though not the letter) of strict layering now appears violated, and the principle is in any case left with very little work to do.

The prevailing approach, however, as represented by Nespor and Vogel (1986), retained a richer conception of the hierarchy, and here strict layering gained general currency, resulting in representations where prosodic trees can always be “sliced” into the various prosodic levels: Levels are never skipped, nor repeated.6 Subsequent analyses, adopting strict layering without discussion, were faced with two chronic problems: “Too little structure” and “too much structure”.

The “too little structure” problem arises because strict layering caps the maximum depth of prosodic trees at n, where n is the number of categories in the prosodic hierarchy. The empirical evidence showed, however, that prosodic structure is more articulate than what can be captured by a closed system of this kind. While it clearly does not have the degree of recursivity seen in syntax, Ladd (1986) presented strong evidence for recursion at the intonational phrase level, arguing that the theory widely overshot its mark by ruling out all recursivity whatsoever (e.g., as in Selkirk 1986: 384: “[P]rosodic constituents of a same category are not nested”).

The “too little structure” problem prompted researchers to posit new categories whenever the prosodic parses required more structure than what was provided by the existing categories. This led to a proliferation of categories (“clitic group”, “intermediate phrase”, “accentual phrase”, “minor phrase”, “major phrase”, “superordinate minor phrase”, etc.; see Ito and Mester 2007b for further discussion). With each new category, any hope of substantiating the hierarchy as a truly universal one receded further into the distance.

Ironically, the strict layering doctrine is also confronted with the opposite problem: “Too much structure”. In order to fulfill strict layering, representations have to be padded to fill every layer, leading to the positing of categories emptied of their intrinsic meaning. This line of argument was laid out in detail in Ito and Mester (2003) with evidence from Japanese showing that a form like anime is structured as \( \omega [o(r\tilde{a}\tilde{a})\tilde{a}] \), with an unfooted last syllable directly dominated by \( \omega \), not as \( [o(r\tilde{a}\tilde{a})(\tau-r\tilde{a})] \), with a degenerate second foot “f”, or as \( [o(f(r\tilde{a}\tilde{a})\tilde{a})] \), with a superordinate “superfoot” F.

4 Leaving the determiner unfooted but parsed as a prosodic word \([o\tilde{a}]\) violates the universal Headedness requirement, and is ruled out as in (16) above.

5 Note that (20d) does not violate Lex-to-ω since the determiner the is a function word.

6 Interestingly, in their earlier work Nespor and Vogel (1983) had still made crucial use of recursive prosodic structures that violated strict layering. As Ladd (1996) points out, strict layering seems to have quickly gained the status of a self-evident truth, so much that it was literally built into the formalization of prosodic structure proposed in Pierrehumbert and Beckman (1988) without so much as reference either to the name of the hypothesis (“strict layering”) or to the work it was first introduced (Selkirk 1984).
Empirical evidence for strict foot-layering would be a language similar to English where words like bo'banana are actually parsed as [\(\bar{\text{o}}\) (ber)](f'\text{nae})\] with a secondary stress (and a full vowel) in the first syllable that is exclusively due to the extra foot required by strict layering. It turned out that this kind of outcome is rarely, if ever, attested, and language after language instead presents us with outcomes like the one in English, well captured by the representation [\(\bar{\text{o}}\) (ber)](f'\text{nae})\] with an unfooted first syllable in violation of strict layering. Positioning a representation like [\(\bar{\text{o}}\) (ber)](f'\text{nae})\], to live up to the facts while upholding the letter of strict layering, is counterproductive because it deprives the category foot of its intrinsic meaning for a language like English, leading to constructs like "stressless foot", "headless foot", etc. In such approaches, strict layering has become a mere convention of how to draw prosodic trees, and few, if any, empirical claims are connected with the constituency posited.

The weak layering theory of prosodic structure of Ito and Mester (2003) expresses this kind of view, and presents it as a response to the "too much structure" problem. "Strict layering" is no longer a monolithic requirement, but is reduced to its more basic components, including constraints enforcing headedness, exhaustive parsing, and left alignment of morphological word to foot. The main interest of this work was on word-internal prosody and prosodic morphology, and the goal was to find a theoretical framework that would make sense of the empirically justified distribution of feet, as described earlier.

This is the approach that was later incorporated into OT. In the formalizations of weakly layered conceptions that appeared in the mid-90's (most prominently Selkirk 1996, a study of function words; see also Peperkamp 1997, among others), other constraints enforcing strict layering were also proposed as violable constraints. In particular, Selkirk proposes a specific constraint (dubbed EXHAUSTIVITY) that enforces strict layering, replicating earlier proposals in pre-OT work.

(22) EXHAUSTIVITY (Selkirk 1996: 190):
No C\(i\) immediately dominates a constituent C\(j\), \(i < j - 1\) [...]

At first glance, this seems to capture the essence of the weak layering conception in a very straightforward way. On second thought, such direct transfers of earlier proposals into OT-phonology are often problematic. Here, it turns out that EXHAUSTIVITY (split into subconstraints such as EXHAUSTIVITY-\(\gamma\), EXHAUSTIVITY-\(\omega\), etc.) is redundant with respect to PARSE-INTO-X. The perspective is certainly different: Whereas PARSE-INTO-X starts with the terminal string and asks whether it has been parsed into all the levels of the prosodic hierarchy, EXHAUSTIVITY looks at the tree structure itself, and scrutinizes the daughters of every node.

However, it is still true that a candidate that entirely fulfills PARSE-INTO-X, for all levels \(X\) of the hierarchy, cannot help also fulfilling EXHAUSTIVITY, showing that a theory that contains both groups of constraints is redundant. We follow the basic approach to parsing theory in OT laid out in Prince and Smolensky (2004) (see also section 2.1 above) and interpret exhaustivity not as a separate constraint, but as an effect of PARSE-INTO-X.

3. Function word complexes in English and German

We turn now to the main topic of this paper, namely, the prosodic form assigned to syntactic structures headed by functional categories as in (23).

(23) a. DP
   D
   \(\bar{\text{N}}\)
   \text{the dinosaurs}
   fnc

   b. IP
   IP
   I
   J
   \text{could live}
   fnc

   c. PP
   PP
   P
   N
   \text{on earth}
   fnc

Such function word complexes, illustrated in (24) for a variety of languages, typically have a characteristic prosodic form where the function word is tightly bound to its lexical host and in some sense subordinated to it.

(24) til Rhodos Danish à Rhodes French
    a Rodi Italian naar Rhodos Dutch
    nach Rhodos German rodosu-e Japanese
    to Rhodes English eic Rodon Greek
    a Rodos Spanish

Generally speaking, prosodic subordination is reflected in the exemption of functional categories and their projections from mapping constraints (see e.g. Truckenbrodt 1999: 226). In our terms, D, I, P, and other functional categories are not subject to LEX-MAPPING (12).7

---

7 A reviewer points out that "function word" may not be a well-defined and independently motivated notion. For example, Wagner (2005) argues that the distribution of function words reflects the fact that such words occur with a high frequency in prosodically subordinated environments and therefore tend to develop reduced allomorphs over time. As we will argue in the next section, even prosodically unreduced function words like suppose\(\alpha\) pattern with function words and not with lexical words, which indicates that the frequency-based explanation might not be sufficient. See
This much is standard practice – but what exactly does it mean for a functional element to be prosodically integrated with its lexical host? Earlier work on the prosodic hierarchy subscribing to strict layering (Hayes 1984, 1989, further developed in Nespor and Vogel 1986) introduced a specific constituent, the “Clitic Group”, to parse exactly such combinations. The Clitic Group as a category different from word and phrase has remained controversial, however, and in line with our general strategy favoring a sparse hierarchy with essential appeal to adjunction and relational structure (Ito and Mester 2007ab), we follow Selkirk (1996) in identifying four possible sites for function words in a prosodic structure without any categories intervening between ω and φ.

(25) Prosodic sites for function words

a. full-ω fncl

b. amalgamated fncl

c. ω-joined fncl

d. φ-attached fncl

These four structures are exactly those predicted to arise from the interaction of the parsing constraints and other relevant constraints discussed in section 1.2.

The (unranked) tableau is repeated here in (26), using a German example of a Det-N sequence with a reduced form of the indefinite determiner (*ne instead of eine), see Hall (1999: 104).

(26) The extended prosodic word

The optimal parsing of fncl-lex structures thus depends on the ranking of the constraints FTBIN, LEX-TO-ω, NO-RECURSION, and PARSE-INTO-ω/φ. Several questions arise at this point. Which structure is correct for a particular language, such as English or German? Is each of the four structures instantiated in some language? In other words, are all four constraints violable? Or are some of them undominated in every grammar and thus potentially assigned to GEN, together with the basic representational constraints in (4)? For example, FTBIN has sometimes been argued to be universally unviolable.

A number of crosslinguistic studies (including Selkirk 1996 for English and Serbo-Croatian and Peperkamp 1997 for various Italian dialects) have addressed these questions, with the goal of matching the empirically attested typology of fncl-lex structures with a factorial typology like the one in (25) (see also the papers collected in Hall and Kleinhenz 1999). Our goal here is the more modest one of (re)asking the question regarding the prosodic structure of such fncl-lex combinations in English and German, two of the more well-discussed cases in the literature.

In an influential paper, Selkirk (1996) opts for the φ-attached structure (25d) for English proclitics, and this position has been adopted for German proclitics in Hall (1999), as well as in a subsequent study by Kabak and Schiering (2006). We will review the previous arguments here, and consider new evidence along the way. From the perspective of the PARSE-INTO-X approach advocated here, it turns out that for both English and German, the overall evidence favors ω-adjunction (25c) over φ-attachment (25d), echoing a line of analysis familiar from earlier work (such as Booij’s (1996) treatment of proclitics in Dutch).
3.1. Between independence and fusion

Past research has uncovered a significant amount of phonotactic and prosodic evidence showing that in many languages (including English and German) function words are neither structured as independent \( w's \) (25a), nor simply prosodically fused with neighboring lexical words into a single \( w \) (25b). Informally speaking, (25a) gives too much independence and weight to the prosodically deficient function word, whereas (25b) gives too little prosodic independence to the lexical word. The structures in (25c) and (25d) avoid these problems: The function word is by itself not given prosodic word status, while the lexical word keeps its prosodic integrity.

The unavailability of fully independent (27a) and fully fused (27b) parses in a language follows from high-ranking FtBin and Lex-to-\( \omega \), respectively: The first constraint militates against establishing a subminimal \( fnc \) as a \( \omega \), the second against turning a lexical word into a prosodic non-entity.

In English (28a) and German (28b) (both standardly analyzed with trochaic foot structure and a right-aligned quantity-sensitive head trochee, see Pater 2000 and Alber 1998), single pretonic syllables remain stressless in initial position.

Foot parsing is non-exhaustive in (28) because Parse-into-f is dominated by FtBin, as shown in (29) (see also section 2.2).\(^9\)

On the other hand, we find secondary stress on the initial syllable in morphologically related forms in (30).

Although morphologically complex, the forms in (30) constitute single prosodic words and are rhythmically different from the function word-lexical word complexes in (32). The latter, but not the former, show initial lapses: A sequence of unstressed syllables constituted by a subminimal function word (the indefinite \( a \) and its German equivalent 'ne) followed by the unfooted initial syllable of the lexical word.

---

\(^9\) In addition, because of a constraint against stress clash.

\(^{10}\) Odd-numbered initial sequences of light syllables receive left-aligned footing (McCarthy and Prince 1993a).
unstressed syllables, violating \textsc{parse-into-f}, as schematically indicated in (35a).
Why does the grammar not instead select the structure \([\langle \text{a} \rangle \langle \text{gu:na} \rangle]\), where the
two syllables are footed? The answer is that the resulting structure (35b)
violates proper bracketing (see (4b) above), a basic requirement on prosodic
structures that, we assume, is properly a part of GEN.

\[
(35) \quad \begin{array}{c l}
\text{a.} & \alpha (\text{or } \varphi) \\
\text{b.} & * \alpha (\text{or } \varphi)
\end{array}
\]

The upshot is that a sequence of unstressed syllables is allowed to arise in
adjuncted structures, but not in prosodically simplex words like (31).

For German, the \(\omega\)-initial \textsc{parse-into-f} effect has an interesting repercussion
in the allomorphy of the past participle of verbs. The past participial form
adds, besides a suffix (-\(e\)) for weak verbs, -\text{en} for strong verbs), an unstressed
prefix ge- (lieben, ge-\text{l}ieb-t, 'love, loved'; stehen, ge-st\text{en}d-\text{en} 'stand, stood'),
a combination sometimes regarded as a "circumfix". The prefix ge-
appears only when the following syllable carries the main word stress (36a), and is absent
otherwise (36b). When the location of main stress varies, the presence/absence of
\text{ge-} varies accordingly (36c).

\[
(36) \quad \begin{array}{c l}
a. & (\text{hat}) \text{ ge-\text{predi}g-t} \\
b. & (\text{hat}) \text{ ge-\text{kib}ebitz-t} \\
c. & (\text{hat}) \text{ ge-\text{schm}\text{a}r-t\text{otz}t}
\end{array}
\]

The extended prosodic word

unstressed syllables, violating \textsc{parse-into-f}, as schematically indicated in (35a).
Why does the grammar not instead select the structure \([\langle \text{a} \rangle \langle \text{gu:na} \rangle]\), where the
two syllables are footed? The answer is that the resulting structure (35b)
violates proper bracketing (see (4b) above), a basic requirement on prosodic
structures that, we assume, is properly a part of GEN.

\[
(35) \quad \begin{array}{c l}
\text{a.} & \alpha (\text{or } \varphi) \\
\text{b.} & * \alpha (\text{or } \varphi)
\end{array}
\]

The upshot is that a sequence of unstressed syllables is allowed to arise in
adjuncted structures, but not in prosodically simplex words like (31).

For German, the \(\omega\)-initial \textsc{parse-into-f} effect has an interesting repercussion
in the allomorphy of the past participle of verbs. The past participial form
adds, besides a suffix (-\(e\)) for weak verbs, -\text{en} for strong verbs), an unstressed
prefix ge- (lieben, ge-\text{l}ieb-t, 'love, loved'; stehen, ge-st\text{en}d-\text{en} 'stand, stood'),
a combination sometimes regarded as a "circumfix". The prefix ge-
appears only when the following syllable carries the main word stress (36a), and is absent
otherwise (36b). When the location of main stress varies, the presence/absence of
\text{ge-} varies accordingly (36c).

\[
(36) \quad \begin{array}{c l}
a. & (\text{hat}) \text{ ge-\text{predi}g-t} \\
b. & (\text{hat}) \text{ ge-\text{kib}ebitz-t} \\
c. & (\text{hat}) \text{ ge-\text{schm}\text{a}r-t\text{otz}t}
\end{array}
\]

The extended prosodic word

unstressed syllables, violating \textsc{parse-into-f}, as schematically indicated in (35a).
Why does the grammar not instead select the structure \([\langle \text{a} \rangle \langle \text{gu:na} \rangle]\), where the
two syllables are footed? The answer is that the resulting structure (35b)
violates proper bracketing (see (4b) above), a basic requirement on prosodic
structures that, we assume, is properly a part of GEN.

\[
(35) \quad \begin{array}{c l}
\text{a.} & \alpha (\text{or } \varphi) \\
\text{b.} & * \alpha (\text{or } \varphi)
\end{array}
\]

The upshot is that a sequence of unstressed syllables is allowed to arise in
adjuncted structures, but not in prosodically simplex words like (31).

For German, the \(\omega\)-initial \textsc{parse-into-f} effect has an interesting repercussion
in the allomorphy of the past participle of verbs. The past participial form
adds, besides a suffix (-\(e\)) for weak verbs, -\text{en} for strong verbs), an unstressed
prefix ge- (lieben, ge-\text{l}ieb-t, 'love, loved'; stehen, ge-st\text{en}d-\text{en} 'stand, stood'),
a combination sometimes regarded as a "circumfix". The prefix ge-
appears only when the following syllable carries the main word stress (36a), and is absent
otherwise (36b). When the location of main stress varies, the presence/absence of
\text{ge-} varies accordingly (36c).

\[
(36) \quad \begin{array}{c l}
a. & (\text{hat}) \text{ ge-\text{predi}g-t} \\
b. & (\text{hat}) \text{ ge-\text{kib}ebitz-t} \\
c. & (\text{hat}) \text{ ge-\text{schm}\text{a}r-t\text{otz}t}
\end{array}
\]

The extended prosodic word

unstressed syllables, violating \textsc{parse-into-f}, as schematically indicated in (35a).
Why does the grammar not instead select the structure \([\langle \text{a} \rangle \langle \text{gu:na} \rangle]\), where the
two syllables are footed? The answer is that the resulting structure (35b)
violates proper bracketing (see (4b) above), a basic requirement on prosodic
structures that, we assume, is properly a part of GEN.
as to allow attachment of the unstressable prefix only to main stress.13 Thus, [ge(∗predignt)] with an initial unstressed syllable is allowed, but [gestu(∗dier)] with two initial unstressed syllables is disallowed. Rather than forcing a stress on the schwa syllable ge by footing it together with the initial syllable of the stem [(ge.stu)(∗dier)], the allomorph without the ge-prefix is chosen [stu(∗dier)].14

Besides the prosodic (rhythmic) evidence showing that every lexical word must initiate a ω, there is also segmental phonotactic evidence in American English against prosodifying the functional complex into a single ω. In identical prosodic environments (stress vs. unstress), the juncture between function word and lexical word acts phonotactically as a ω-juncture. In fnc-lex combinations (a [r]Annie’s) and lex-lex combinations (ea[r]apples), onsets of stressed syllables show flapped [r]. This is in sharp contrast with the word-internal situation, where such flapping is impossible (a[r]ack). Similarly in unstressed environments, clearly ω-internal [t] is flapped (pi[r]y, capi[r]alism, see Steriade 2000 and Davis 2005 for different analyses of the additional paradigm uniformity factors at work here) but lex-initial [t] is aspirated in fnc-lex combinations (a [t]omato) just as in lex-lex combinations (grow [t]omatoes).

---

13 The ge-rule extends to the nominal ge-prefix in pejorative nouns (Ge-sing-e ‘singing’, Getanz-e ‘dancing’, etc.), which is significant because the latter can be attached quite productively in colloquial speech. Harden (2003) shows that speakers’ judgments when presented with nonce-forms where the prefix is not followed by main stress essentially follow the rule, with some interesting variation in judgments between fully unacceptable and marginal forms (*Ge-be-wirt-e ‘hosting’, *Ge-neutrale-e ‘calling’, *Ge-studier-e ‘studying’). Overall, the very existence of the ge-rule is an indication that the schwa-containing prefixes ge- and be-, different from prosodically non-cohering unstressed prefixes like ent- and ver- (which do not allow resyllabification), are prosodically integrated with the stem in a simplex ω, and are not part of a recursive structure (such as argued by Booij 1996 for Dutch). Otherwise (cf. (29)), there is no systematic reason why ge- should be absent in (36b).

14 Some examples provided by a reviewer with a sequence of otherwise unstressed verbal prefixes (e.g., ver-) show the robustness of Parse-Into-f in prosodically simplex words. In order to avoid stressing the ver-prefix, the past participle forms of ver’schafft ‘seize’ and ver’kaufen ‘sell’ also choose the ge-less allomorph: ver’schafft, not *(versch)(ge)(∗kaft), and ver’kaufen, not *(ver)(ge)(∗kaft). However, when the stem itself begins with a stressless syllable (Studier’ ‘study’) or with a stressless prefix (Be’antem ‘a civil servant’), ver- can receive secondary stress, as in (verstud’ier) ‘mis-studied’, (verkehr’antem) ‘make into a civil servant’. In this case, there is no stem allomorph that improves the prosody: *ver’dier, *ver’antem. Similarly, ungläublich ‘unbelievable’ but with secondary stress in unver’antwortalich.

15 See Ito and Mester (2007) for further discussion regarding the word-onset requirement forcing ambisyllabicity.

---

The syllabification patterns of German also argue against full ω-integration of the functional word complex. Contrary to what full prosodic fusion would predict, there is no re- or ambi-syllabification to onset in fnc-lex complexes ([ein.Au.to ‘a car’, ein.A’the.ner ‘an Athenian’, *ei.nAu.to. *ei.nA.the.ner], just as in lex-lex combinations ([kauf Auto.to ‘buys cars’, *kauf.Au.to.to]). This even holds for function words reduced to a single consonant, such as ‘n for ein: ’n Auto and ’n Athenner have syllabic nasals connected to the onset of the following syllable: [E)nAu.to., not *[E.n]Au.to. or *[E.n]Au.to.16

Overall, the case against fully independent [ωfnc,ωlex] and fully incorporated [ωfnc,ωlex] parses of function word complexes is strong because of the variety of counter evidence, including both prosodic and phonotactic facts and generalizations. On the other hand, all the evidence considered is compatible with structures of the general form [ωfnc,ωlex], whether x=ω (ω-adjoined fnc) or x=φ (φ-attached fnc), in both of which the lexical word has prosodic word status but the function word does not.

We are left, then, with the choice between the two structures repeated here in (37). As indicated, the difference is whether Parse-Into-ω or No-Recursion is violated. Convincing evidence for either choice is harder to come by, and many researchers have simply assumed one or the other structure. The ω-adjoined structure has been explicitly argued for by Booij (1996) for Dutch, Vigário (1999) for European Portuguese, and Ito and Mester (to appear) for English, among others. On the other hand, Selkirk (1996) and Hall (1999) have presented arguments for φ-attachment of function words (in English and German, respectively).

(37) a. ω-adjoined fnc

b. φ-attached fnc

---

16 It is true that in Standard German lack of resyllabification goes hand-in-hand with systematic [?] insertion at the beginning of vowel-initial words (ein.[?]Au.to., etc.), but it would be wrong to conclude from this that everything can be reduced to the phonetic presence or absence of a glottal word onset. Among dialects without systematic [?] insertion, some allow resyllabification (Swiss German) and some do not (Bavarian German, see also Gutch 1992), indicating that this is a separate parameter of variation (Kabak and Schiering 2006: 65).
Before laying out our case for \(\omega\)-adjunction, we first review the evidence that has been presented for \(\varphi\)-attachment, which will turn out to be inconclusive.

3.2. \(\varphi\)-attachment: reassessing the arguments

The evidence presented in Selkirk (1996) and Hall (1999) in favor of \(\varphi\)-attachment has a negative form: Instead of positively identifying some property of \(x\) in \(\lfloor fnc, lex \rfloor\) that clearly belongs to \(\varphi\) and not to \(\omega\), hence \(x = \varphi, x\) is claimed to lack a certain \(\omega\)-property and is for this reason alone set equal to \(\varphi\) – faute de mieux, so to speak. Selkirk (1996: 197-198) argues as follows: In \(\omega\)-initial position, voiceless stops are aspirated even when the first syllable is unstressed and not foot-initial: 

\[\text{grow t}^{b}\text{omatoeS, grow t}^{b}\text{hétunias, grow t}^{b}\text{älendula}.\]

Recall that this was one of the phonotactic pieces of evidence in favor of the \(\omega\)-hood of \(\text{lex}\) in \(\text{fnc-lex}\) complexes \(\lfloor a\lfloor t^{b}\text{omato}\rfloor\rfloor\), which both of the structures in (38) predict. However, in terms of the aspiration constraint on \(\omega\), the two structures make different predictions regarding the initial aspiration of the larger constituent.

\[(38)\]

<table>
<thead>
<tr>
<th></th>
<th>(\omega)-adjointed fnc</th>
<th>(\varphi)-attached fnc</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f)</td>
<td>aspiration predicted</td>
<td>aspiration predicted</td>
</tr>
</tbody>
</table>

Pointing to examples as in (39), where the \(fnc\)-initial \(r\)'s can remain unaspirated (and be realized as flaps instead), Selkirk (1996) concludes that \(\omega\)-adjunction (38a) makes the wrong prediction and that (38b) is correct, where the function word is not \(\omega\)-parsed and immediately attached to \(\varphi\).

\[(39)\]

a. Take Grey \([t]o\) London.
b. They grow \([t]o\) the sky.
c. We're going to fly \([t]o\) Seattle on Monday.

The argument here implicitly makes a specific claim about the way domain-initial fortition works in the prosodic hierarchy: The left boundary of \(\omega\) is a position of fortition, but not the left boundary of \(\varphi\). This claim encounters some problems. First, we find aspiration marking at the left boundaries not only of \(t^{17}\) and \(\omega\), but also of \(f\) and \(v\) – in particular, as (40) shows, when a functional complex occupies initial position.

\[(40)\]

a. \([t]o\) London. \([t^{b}]o\) London. \([t^{b}]o\) Seattle. 
b. \([t]o\) to London. \([t^{b}]o\) Seattle. 

Taken at face value, this leaves us with the overall picture in (41), where all categories from \(f\) upwards mark their left edges by aspiration – except for \(\varphi\).

\[(41)\]

<table>
<thead>
<tr>
<th>category</th>
<th>aspiration at left edge?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f)</td>
<td>yes</td>
</tr>
<tr>
<td>(\omega)</td>
<td>yes</td>
</tr>
<tr>
<td>(\varphi)</td>
<td>no</td>
</tr>
<tr>
<td>(v)</td>
<td>yes</td>
</tr>
</tbody>
</table>

This should give us reason to pause: The observations in (39) are not in doubt, but do they actually prove that, of all left edges in the hierarchy, \([\varphi-]\) is not a position of fortition?

Category-by-category stipulations of specific processes were a hallmark of the early days of prosodic hierarchy theory, when they were taken as a positive result since they seemed to provide evidence for the existence of the hierarchy in the first place. In this case, however, even if one accepted the claim at face value, one would, at a minimum, still wonder why it is precisely \(\varphi\) that is the odd one out.

At the other extreme, some might take the picture in (41) as a good reason to doubt the very existence of \(\varphi\). We remain persuaded by the overwhelming evidence for a level of phrasing corresponding to \(\varphi\), and instead argue that \([\varphi-]\) is in fact a position of fortition, just like the other categories in (41), and that something else is going on in (39).

In a general vein, this is in line with recent research which has accumulated a significant amount of evidence (see Wagner 2005: 131–136 for an overview) showing that the differences between prosodic boundaries at the various levels of the hierarchy do not in general concern the type of effect but rather its strength, which is hierarchically cumulative, as a broad generalization. Category-specific stipulations, such as the putative exemption for \(\varphi\) in (41), have no place in this

\[17\text{ Or alternatively, \(\text{"stressed }\varphi\).} \]
much more restrictive conception of the effects that are associated with the hierarchy. For the case in point, fortition requirements such as aspiration are not distributed in an on-again, off-again fashion over the prosodic hierarchy, with each category being free to make its own stipulations independent of other categories. Rather, experimental work has demonstrated hierarchical cumulativity. Thus, Fougeron and Keating (1997) conducted a systematic study of articulatory strengthening in American English at different levels of the prosodic hierarchy. Using reiterant speech, with each syllable of a model sentence replaced by [no], linguopalatal contact was measured for consonants and vowels initially, medially, and finally in \( \omega, \phi, \iota \), and \( \upsilon \). The general finding was that “the higher the prosodic domain, the more linguopalatal contact the consonant has” (Fougeron and Keating 1997: 3728). A later study (Cho and Keating 2001) replicates and extends these results for Korean, and Flack (2006) collects a large amount of cross-linguistic evidence showing that fortition requirements are upwardly uniform across word boundaries (as in compounds, cf. ho[t] [\( \text{\textipa{tub}} \)], but not \( \omega \)-internally, and are one of the indicators for \( \omega \)-junctures.

In looking for a solution, let us consider the underpinnings of the phonotactic generalizations. Structurally speaking, aspiration is not a hallmark of \( \omega \)-initial segments per se, but occurs when it clearly marks the beginning of a structural unit, here \( \omega \). In terms of Ito and Mester (1999), it is a crisp-edge property. In both structures in (43), \( y \) is \( \omega \)-initial: \( \omega \)-initial segments are aspirated at crisp (\( \omega \))-edges (43a), and flapped when there is no crisp edge (43b), i.e., when the segment is ambisyllabic (Kahn 1976).

\[
\begin{align*}
\text{(43) a. crisp edge} \\
\text{b. non-crisp edge}
\end{align*}
\]

The crisp-edge structure also accounts for the doubled [t] pronunciation in I forgo[t] [\( \text{\textipa{tlo call}} \)]. Clusters of adjacent identical consonants in English arise across word boundaries (as in compounds, cf. ho[t] [\( \text{\textipa{tub}} \)], but not \( \omega \)-internally, and are one of the indicators for \( \omega \)-junctures.

What is special about the non-crisp cases under discussion, then, is not just the phonetic fact that the segments are flapped, but rather that (onset-)ambisyllabification has here applied to a limited set of \( \omega \)-words.\(^{20}\) Goldsmith (2001) observes that the to in fly to Seattle (39c) has a close phonological relationship with the word to its left. There is no left-alignment constraint for function words

\[^{18}\text{Tomato is a member of the Oxford 3000 list of the 3000 most frequent words of the English language (see http://www.oup.com/elt/catalogue/teachersites/oald7/oxford.3000/oxford.3000_list?cc=gb), but still never lenites to \( *[t] \text{omato} \) in any context.}\]

\[^{19}\text{Goldsmith (2001) notes the interesting contrast between I've go[r] to leave soon vs. I forgo[t] [\( \text{\textipa{tlo leave}} \).}\]

\[^{20}\text{A reviewer points out that this option might arise out of the availability of reduced allomorphs for these items, which tend to be deaccented and attached or "encliticized" to the preceding intonational phrase when occurring sentence-finally. The combination of being frequent and occurring in this "weak" prosodically subordinated environment might make a development of a reduced allomorph likely. This line of thought is an interesting alternative to the standard approach based on the notion "function word".}\]
(no FNC-TO-ω constraint), and contraction facts are also indicative of such a close relationship. Besides the celebrated wanna contraction, we find lexicalized enclitization in cases like gotta (got to), gonna (going to), supposed (supposed to), haflia (have to), and oughta (ought to). The cases in (42) show lexical words (adverbs) showing a similar kind of behavior, which might be related to the fact that they are highly predictable in such conversational routines.

While further investigation is needed to determine the exact environment of ambi syllabification, and a unified explanation of to-words and other details await further phonetic studies, we here conclude that the flapping evidence does not constitute a strong argument for ω-attachment of fnc. In fact, some of the crisp-edge evidence might be construed in favor of ω-adjunction. The argumentation for ω-attachment in German in Hall (1999: 124–126) proceeds along lines similar to Selkirk's (1996). To account for the limited distribution of initial ω-syllables, Hall posits two constraints, one (Hall 1999: 114) barring schwas from occurring ω-initially (*[ωθ . . .]), and another one (Hall 1999: 126) “that ensures that pwords do not begin with a consonant other than [b] or [g] plus schwa” (e.g., B[α]ginn ‘beginning’ and g[α]nau ‘exactly’).

The constraint specifically barring w-adjunction in German in Hall (1999: 124–126) proceeds along lines similar to Selkirk's (1996). To account for the limited distribution of initial ω-syllables, Hall posits two constraints, one (Hall 1999: 114) barring schwas from occurring ω-initially (*[ωθ . . .]), and another one (Hall 1999: 126) “that ensures that pwords do not begin with a consonant other than [b] or [g] plus schwa” (e.g., B[α]ginn ‘beginning’ and g[α]nau ‘exactly’).

(44) a. [β]s kommt ‘it comes’
   b. ts[a] Hause ‘at home’
   c. ’n[α] Machine ‘a machine’

Citing forms such as those in (44), Hall points out that the ω-adjoined [ωnfnc[ωlex]] structure (as well as the fused [ωnlclex] structure) would not fulfill the posited schwa requirements on prosodic words: (44a) violates the ω-initial schwa constraint, and (44b,c) the constraint barring ω-initial Cω except for bo

and ga. The fnc must, so the argument goes, be directly attached to ω, where these schwa-related requirements hold.

The strength of the argument here depends on the analysis on which it rests, in particular, on the constraints that drive it. The constraint specifically barring ω-initial Cω, with a built-in exemption of the elements bo and ga, is awkward as a phonological markedness constraint since it is unlikely to find cross-linguistic support. In German, it encodes the morphological fact, arbitrary in terms of phonology, that bo- and ga- but not, for example, ko-, do-, no-, or lo-, happen to exist as prefixes. The restriction should therefore be regarded as a distributional generalization on lexical words, rather than a requirement on the shape of prosodic words. A possible formal analysis might invoke constraints on prominent positions, as developed in the work of Smith (2004), where the initial position of content words plays a central role, reflecting its psycholinguistic importance. Functional material, irrespective of its prosodic parsing, is not prominent in this way, and hence not subject to such constraints. The explanation therefore does not depend on whether the fnc is ω-adjoined or ω-attached.

In sum, both Selkirk (1996) and Hall (1999) provide negative evidence against ω-adjunction rather than positive evidence for ω-attachment. There is no intrinsic reason, connected with known and recognizable properties of ω, why a proclitic fnc such as to must be ω-attached. Rather, some property associated with lex in ω-initial position is absent in proclitic fnc, and relegating it to ω seems to remove the problem. In German, this is not the case however, once we look at the properties of ω in the context of the whole hierarchy. The solution, we have argued, is to move away from constraints that directly encode some descriptive generalization about ω (such as beginning with at most one unstressed syllable), which are useful as preliminary diagnostics for investigation, but not genuine elements of the theory. Explanations should rather be sought in the interaction of several general, well-motivated, and independent constraints, such as LEX-TO-ω, FTBIN and PARSE-INTO-f, two of which do not even mention ω.

For the cases under discussion, we have suggested that the non-crisp alignment of functional material makes a difference at ω-edges, as well as constraints on prominent positions, which already incorporate a psycholinguistically grounded distinction between content words and function words. On this basis, we can understand in a more principled way why function words and content words can both occupy ω-initial position and still differ in some respects, and the evidence is fully compatible with the ω-adjoined structure.

---

21 See Jurafsky, Bell, Gregory, and Raymond (2000) on the role of frequency and predictability from neighboring words in reduction processes like flapping.

22 There appear to be dialectal differences, and some speakers do not produce onsetless schwa-syllables in phrase-initial position. For such speakers, *[β]s kommt ‘it comes’ is not a possible pronunciation even in a speech register where other fnc reductions take place: *[β]*s aβ n[α] Banane ‘it ate a banana’. The dialectal difference cannot be attributed to whether fnc is ω-adjoined or ω-attached (as argued by Selkirk 1996 for dialects of Serbo-Croatian, and by Peperkamp 1996 for dialects of Italian), since Hall’s (1999) schwa-requirements would favor the ω-adjoined structure for [ωas [ωaβ]] and the ω-attached structure for [ωaβω Banane]. Such a situation is also quite problematic for the constraints determining the prosodic structuring of fnc-lex complexes: The ranking of PARSE-INTO-ω and NO-RECURSION would have to vary from function word to function word, depending on a particular segmental property (V- or C-initial), a situation not otherwise encountered.
4. Evidence for ω-adjunction

In this section, we will lay out our case for ω-adjointed function words. We first show that the special prosody of ϕ-final function words provides one of the key arguments in favor of Strict Layering (Selkirk 1984: 366-382). The generalization to be captured is illustrated in (45): At the right end of a phrase, a monosyllabic function word must appear in its strong form, not in its destressed form.23

(45) Phrase-final
a. They boxed the crowd [m].
   They fought [m]/[on] the ring.
b. Did you let the cat [m]?
   Would you sit [m]/[on] the car?
c. I'd like [m], but I'm not sure
   I [kan]/[f][ud].
d. What were you thinking [av]
   I was thinking [av] Paris.
e. What did you look [æt] yesterday?
   Take a look [æt] Tiffany's.

No special provision for phrase-final function words is needed, so the argument goes, because their promotion to the full status of an independent prosodic word follows directly from Strict Layering (Selkirk 1984: 366; Selkirk and Shen 1990: 332-335; Selkirk and Tateishi 1988).

(46) *(ϕ, What did you look at) *(ϕ, yesterday)?
   *(ϕ, [ϕ, look at]) *(ϕ, [ϕ, yesterday])
   *(ϕ, [ϕ, look]) *(ϕ, [ϕ, yesterday])
   *(ϕ, [ϕ, look] at) *(ϕ, [ϕ, yesterday])
   *(ϕ, [ϕ, look] at) *(ϕ, [ϕ, yesterday])

The general proclisis pattern of English means that fnc cannot cliticize to the left, making (46a) impossible,24 and proper bracketing prohibits cliticizing across ϕ.

23 This generalization holds also when the function word is a phrase by itself, where it also occupies the right end of a phrase.
24 As is well known, the overwhelming default in English is proclisis, not enclisis, in line with the general syntactic pattern of the language. Pronominal object enclitic forms like gimme, got'm, or need'm go against this general proclisis imperative and have a special morphosyntactic status (Selkirk 1984: 383-406, 1996: 459-460). We return to this point in section 4.3.
25 While phrase-final lengthening is clearly a widely attested phenomenon, as a reviewer reminds us, it should not be confused with prominence or strengthening. On the other
an explanation, preserving a central insight of Strict Layering Theory. On the other hand, direct \( \varphi \)-attachment of function words as a default, by passing the prosodic word level, means that this result is lost.

4.2. The argument from structural binarity

Even though less prominent than in syntax, structural binarity considerations play an important role in phonology, schematized in Ito and Mester (2007b) as in (48).

(48) \( \text{BIN}(x) \), where \( x \) is an element of the prosodic hierarchy

\( \text{BIN}(\varphi) \) is the familiar \( \text{FTBIN} \) constraint, and \( \text{BIN}(\sigma) \) the constraint against super-heavy trimoraic syllables. Versions of \( \text{BIN}(\omega) \) and \( \text{BIN}(\sigma) \) have been recruited to account for, e.g., truncation patterns and phonological phrasing (Ito and Mester 2003, Kubozono 1995, Selkirk 2000, etc.). Of relevance are binarity factors related to the parsing of function words discussed by Selkirk (2000) in accounting for the possible variations in prosodic phrasing, as in (49).

(49) She loaned her rollerblades to Robin.
   a. \( \rightarrow (\varphi)(\varphi) \)
   b. \( \rightarrow (\varphi) \)

According to Selkirk (2000), phrasal variation in a neutral non-focused context is permitted in (49), because neither phrasal choice fulfills strict binarity ((49a) has a unary \( \varphi \), (49b) a ternary \( \varphi \)). On the other hand, the candidate parse (50a) fulfills \( \varphi \)-binarity exactly and therefore forestalls the co-selection of the non-binary variant (50b). The structure in (51) illustrates how (50a) fulfills \( \varphi \)-binarity (each \( \varphi \) is binary).

(50) She loaned her rollerblades to Robin’s sister.
   a. \( \rightarrow (\varphi)(\varphi) \)
   b. \( \rightarrow (\varphi) \)

Could one say that it is all a matter of definition? For example, the constraint could be formulated as in (53), and binarity would be fulfilled in (52), despite appearances.

(53) \( \text{BIN}(\varphi) \) (rejected formulation): \( \varphi \) requires exactly two \( \omega \).

This formulation tells us to only look at \( \omega \)-daughters of \( \varphi \) and to ignore everything else, so that branches leading to other daughter nodes simply do not count. But just stating the constraint reveals its artificiality – the necessity to have recourse to a complication like (53) is a counterargument in itself. More importantly, the constraint has the flaw that it repeats a section of the prosodic hierarchy. For example, it is just an accident that it mentions \( \varphi \) and \( \omega \), and not \( \omega \) and \( \sigma \). The upshot is that the simple everyday conception of binarity based on branching works fine – provided the prosodic structure is understood as \( \omega \)-adjunction.

A different way out, pursued in Selkirk (2000: 244), is to postulate an additional prosodic category, \( \text{MiP} \) (minor phrase) composed of \( \text{fnc-lex} \) combinations and intervening between \( \omega \) and \( \varphi \) (= major phrase), and to formulate the binarity constraint as \( \text{BIN}(\text{MiP}) \) “a major phrase consists of just two minor/accentsual phrases”. As illustrated in (54), this makes the overall prosodic form isomorphic to the \( \omega \)-adjunction structure (51).
Occam’s razor, however, militates against such additional categories as long as the existing ones are sufficient to represent the prosodic groupings manifest in the data, and no solid independent evidence for the new categories is available. Positing the minor phrase raises the additional question of how it would differ from the Clitic Group, which has been argued against as a separate category in earlier work (see the beginning of section 3 above).

Whatever the final resolution of these questions is, it remains true that the binarity facts here follow immediately from the ω-adjunction view of function words, whereas the φ-attachment view needs to introduce an additional phrasal category to specifically create prosodic forms isomorphic to the structures that are the hallmark of ω-adjunction.

4.3. The argument from maximal ω-projections

What other factors, then, might be brought to bear on the decision between the two possible structures for fnc-lex complexes in (55)?

(55) a. ω-adjunction: b. φ-attachment:

\[
\begin{array}{c}
\text{MiP} \\
\text{MiP} \\
\text{MiP} \\
\text{MiP}
\end{array}
\]

\[
\begin{array}{c}
\text{MiP} \\
\text{MiP}
\end{array}
\]

\[
\begin{array}{c}
\text{fnc} \\
\text{lex}
\end{array}
\]

\[
\begin{array}{c}
\text{fnc} \\
\text{lex}
\end{array}
\]

Since the two structures have in common that the lexical word is a ω (Lexto-ω is fulfilled), the difference might seem to solely lie in the placement and status of the function word. Indeed the arguments that we have seen above all focus on this point (e.g., Is fnc ω-initial? Is it dominated by ω?). However, another important difference between the two structures lies in the prosodic status not of the function word to but of the lexical word Rhodes: It is a full and independent prosodic word in (55b) [\text{to } [\omega \text{Rhodes}]] but not in (55a), where it is only a segment of the larger prosodic word [\text{to } [\omega \text{Rhodes}]]. In terms of the approach developed in earlier work (Ito and Mester 2007b, to appear), [\omega \text{Rhodes}] is a maximal projection of ω in (55b), but not in (55a), where it is dominated by a higher ω. This makes a difference to the extent that there are phonological processes that are specific to maximal prosodic words. Such processes indeed exist. An example is linking and intrusive r in non-rhotic varieties of English, whose natural interpretation, pace McCarthy (1993), as hiatus breaker / onset filler can only be maintained if maximal and non-maximal ω are distinguished. After a brief review of the basic analysis of English r-sandhi in Ito and Mester (to appear), we pursue the consequences of such an approach in terms of the adjunction status of various types of function words. Comparison with German (Kabak and Schiering 2006) provides further arguments for ω-adjoined structures, and thus offers evidence for high-ranking Parse-Into-ω in these languages.

The literature on English r-sandhi (r-loss and r-insertion at certain morphological and syntactic junctures) is extensive, both in descriptive and theoretical terms (see McCarthy 1991, 1993, as well as McMahon 2000, and works cited there). The hiatus-breaking r- occurs after the non-high vowels [a, ɔ, ɑ], as in the idea-r-is this or law-r-and order, and is productively inserted, as Wells (1982: 226) points out, in acronyms (as far as NAFTA-r-is concerned), loanwords (schwa-r-insertion, Lufthansa-r-officials) and even in r-intruders’ pronunciation of foreign languages (ich bin ja-r-auch fertig, j’étais déjà-r-ici, Gloria-r-in excelsis Deo, viva-r-Espaia, etc.). The phenomenon is widespread, and a virtually identical process exists, for example, in Bavarian German (wie-r-ich gesagt habe [wiːˈriː gsgd hab] ‘as I said’).

An argument that the ω-adjoined structure has advantages over φ-attachment comes from an at first glance puzzling restriction on the process, which is otherwise fully automatic and productive: Intrusive -r- does not appear after function words. This was first noted in Kahn (1976), and McCarthy (1993), the source of the examples in (56), provides an analysis (see also McCarthy 1991).

(56) a. Why do Albert [do ælбот] and you *[tə ælбот] b. quarter φ(t) eight [tə ət] *[tə ət] c. the apples [tə æp] *[tə æp] d. to add [tə æd] to (h) is [tə ɪz] troubles, *[tə ɪd], *[tə ɪz]

In (56d), we find unresolved vowel hiatus both in the familiar fnc-lex complex (to add) and in the fnc-fnc sequence (to (h)is). The summary table in (57) gives an overview of the presence and absence of the hiatus breaker with various combinations of lex and fnc.

\[
\begin{array}{c|cc}
\text{lex} & \text{fnc} & \text{to} \\
\hline
\text{fnc} & \text{lex} & \\hline
\end{array}
\]
we show that reference to maximal and minimal projections of \( \omega \) and \( \varphi \) is crucial in understanding the phonological behavior of complex compounds in Japanese (\( \omega_{\text{min}} \) as domain of rendaku, \( \omega_{\text{max}} \) as domain of junctural accent, \( \varphi_{\text{min}} \) as domain of deaccentuation, etc.).

In the context of English function words and intrusive \(-r-\), reference to \( \omega_{\text{max}} \) plays an important role in two situations, viz., for \( \omega_{\text{max}} \) versions of lexical word alignment (60a) and of the onset requirement (60b).

\[(60) \quad \begin{align*}
\text{a.} & \quad \text{LEX-TO-} \omega_{\text{max}} \\
\text{b.} & \quad \text{ONSET-} \omega_{\text{max}}
\end{align*}\]

LEX-TO-\( \omega_{\text{max}} \) (60a), a more specific version of the general alignment constraint LEX-TO-\( \omega \) discussed above in section 2.2, requires alignment of \( \text{lex} \) to a maximal \( \omega \). As shown in (61), ranking the right-handed version of the constraint over the left-handed version favors proclisis, as in English, and the opposite ranking yields enclisis.

\[
\text{LEX-TO-} \omega_{\text{max}} \text{ (60a)}, \text{ a more specific version of the general alignment constraint LEX-TO-} \omega \text{ discussed above in section 2.2, requires alignment of } \text{lex} \text{ to a maximal } \omega. \text{ As shown in (61), ranking the right-handed version of the constraint over the left-handed version favors proclisis, as in English, and the opposite ranking yields enclisis.}
\]

\[
\text{LEX-TO-} \omega_{\text{max}} \text{ (60a)}, \text{ a more specific version of the general alignment constraint LEX-TO-} \omega \text{ discussed above in section 2.2, requires alignment of } \text{lex} \text{ to a maximal } \omega. \text{ As shown in (61), ranking the right-handed version of the constraint over the left-handed version favors proclisis, as in English, and the opposite ranking yields enclisis.}
\]

\[
\text{LEX-TO-} \omega_{\text{max}} \text{ (60a)}, \text{ a more specific version of the general alignment constraint LEX-TO-} \omega \text{ discussed above in section 2.2, requires alignment of } \text{lex} \text{ to a maximal } \omega. \text{ As shown in (61), ranking the right-handed version of the constraint over the left-handed version favors proclisis, as in English, and the opposite ranking yields enclisis.}
\]

\[
\text{LEX-TO-} \omega_{\text{max}} \text{ (60a)}, \text{ a more specific version of the general alignment constraint LEX-TO-} \omega \text{ discussed above in section 2.2, requires alignment of } \text{lex} \text{ to a maximal } \omega. \text{ As shown in (61), ranking the right-handed version of the constraint over the left-handed version favors proclisis, as in English, and the opposite ranking yields enclisis.}
\]
A second important role of the maximal prosodic word lies in the fact that its onset is a position of particular prominence. This is the driving idea of the analysis of English -r- in Ito and Mester (to appear): The appearance of both linking and intrusive r in external sandhi is a manifestation of the requirement for maximal prosodic words to have an onset. The process results from the interaction of two constraints, each a positionally restricted version of a more general one: ONSET-ωmax (62a), a special version of ONSET (62c) for a prominent position. Hall (1999) makes a similar point for German.

(62) Constraints driving r-intrusion

| a. proclisis | ONSET-ωmax, DEP-INIT-ω | lex, c, lex, c, lex, c, lex, c |
| b. enclisis | ONSET-ωmax, DEP-INIT-ω | lex, c, lex, c, lex, c, lex, c |

Ranked as ONSET-ωmax = DEP-INIT-ω, the interaction of these two constraints derives the correct distribution of intrusive -r-, as shown in (63). We find -r- in lex-lex and lex-fnc configurations (63a,c) at the beginning of a maximal prosodic word because of the action of the dominant constraint ONSET-ωmax.33 On the other hand, in fnc-lex and fnc-fnc configurations (63b,d) we are dealing with the onset of a non-maximal prosodic word, where this constraint is irrelevant. Subordinate DEP-INIT-ω therefore has a chance to assert itself, preventing insertion, and the hiatus is not resolved by -r-.34

33 A special case of lex-fnc (see note 24 above) is found in pronominal object enclitic forms like gimme, got'm, or need'm, which, we assume, following previous researchers, have a special morphosyntactic status. The host verb and the enclitic object pronoun together form a single lexical word, which is in turn mapped by Lex-ω alignment into a single prosodic word. Intrusive -r- is found in this context ([iɔdrɔʊr]/[iɔsawʁi], [iɔsubpʰeənaɪt(h)ɪm]) because the pronoun is not ω-initial, and hence not subject to DEP-ω-INIT. The general constraint ONSET is responsible for r-insertion, just as it is in other ω-contexts, such as [iɔsubpʰeənaɪt], [iɔwithdrawal], etc. In a detailed analysis of Dutch clitics, Booij (1996) provides a variety of arguments that the prosodic structure for host-enclitic sequences is different from proclitic-host sequences. Just as in English, the former is prosodically integrated into a single prosodic word, whereas the latter forms an adjacency structure. Hall (1999) makes a similar point for German.

34 Even though intrusive -r- is impossible in fnc-lex complexes, underlying r is obligatory (McCarthy 1993: our answer were eating, for eating, either apples or oranges, for any reason, under any circumstances). The distinction is therefore still part of the synchronic grammar (see Ito and Mester to appear for discussion) and creates a serious obstacle for any attempt to reduce the whole r-sandhi phenomenon to allomorphy, i.e., by listing variants with and without final -r- for all relevant lexical items (e.g., were eating, etc). The allomorphy approach has a basic problem with the productivity and rule-governed character of the phenomenon. Why would all relevant items have a variant with -r? Why would intrusive -r- appear in new loanwords, and in the pronunciation of foreign languages? etc.) and cannot even deal with the basic fnc-lex distinction. One might think that gonnafnc, different from wannaLex and from underfnc, simply lacks a variant with -r. But this does not explain why gonna is suddenly able to sponsor intrusive -r- when it stands in phrase-final position (I said I was gonna, I was gonna, I did, etc., see McCarthy 1993 and section 4.1). One could resort to further listing and add on a "precompiled" phrase-final variant ([iɡəʊtʊl, iɡoʊnæt, iɡəʊtʊl], but this would mean giving up on any attempt to explain the generality of the phenomenon.
The crucial prerequisite for this explanation of the lack of r-insertion after a function word, as in to Anne (63b), is the different prosodic status of Anne in this position — as a subpart of a maximal \( \omega \), not a maximal \( \omega \) all by itself as in (63a), a difference directly due to the \( \omega \)-adjoined structure for function words. Direct \( \varphi \)-attachment (64b), on the other hand, by having the function word skip \( \omega \)-parsing entirely, turns both Anne's into maximal \( \omega \)'s and misses the crucial contrast.

The extended prosodic word

The ONSET-based account of r-intrusion in (62) and (63) rests on the optimality of \( \omega \)-adjoined structures for English function word complexes, which is in turn explained by the ranking Parse-into-\( \omega \) \( \gg \) No-Recursion (65), in terms of the prosodic constraints introduced in section 2.

Selkirk (1996), on the other hand, follows McCarthy (1993) in ascribing r-intrusion to a FINAL-C constraint, which requires every \( \omega \) to end in a consonant. In such an approach, the dominating nodes do not make a difference:

Whether \( fnc \)-lex structures are \( \omega \)-adjoined (66a) (McCarthy 1993) or \( \varphi \)-attached (66b) (Selkirk 1996), \( fnc \) is not placed in a \( \omega \)-final position, hence FINAL-C does not require the presence of -r-. Although successful in capturing the absence of \( r \) in these positions, FINAL-C, in requiring \( \omega \)-final codas, directly conflicts with the universal markedness constraint disprefering codas in general. As shown above in (62)–(63), once the onset requirements for different levels of the prosodic hierarchy are properly understood, in conjunction with appropriately assigned \( \omega \)-adjoined prosodic structures (65), there is no need to appeal to the intrinsically problematic FINAL-C constraint (for further discussion and motivation, see Ito and Mester to appear).

Looking beyond FINAL-C's specific problems, the next section shows that any analysis relying on constraints on \( \omega \)-endings instead of \( \omega \)-onsets, with an assumed across-the-board absence of \( \omega \)-boundaries after function words, becomes untenable when we consider the full spectrum of function word complexes, including prosodically more independent function words such as disyllabic gonna or trisyllabic supposeta.

4.4. The argument from longer function words

Previous research on the status of function words has mostly been concerned with the monosyllabic case, prototypically represented by lone light syllables.
such as to or the which are intrinsically prosodically undersized, unfootable by themselves, often reduced, and unable to constitute an utterance. There is usually little detailed discussion of function words that are large enough to be prosodically full-sized (e.g., disyllabic feet such as after or during), and how they are to be incorporated into the overall prosodic structure.

Although often relegated to footnotes or silently omitted from discussion altogether, such prosodically full-sized function words are neither rare nor atypical, as can be gleaned from the illustrative lists of monosyllabic, disyllabic, and polysyllabic prepositions in English (67) and German (68). Note that even the monosyllables include heavy (footed) items such as down and through.

(67) English prepositions
a. monosyllabic: at, by, down, for, from, in, like, of, on, out, round, since, through, till, to, up, with
b. disyllabic: about, above, across, after, against, along, among, around, before, behind, below, beneath, between, beyond, during, except, over, towards, under, until, within, without
c. polysyllabic: underneath

(68) German prepositions
a. monosyllabic: ab, an, auf, aus, bei, bis, durch, für, in, mit, nach, seit, statt, trotz, um, von, vor, zu
b. disyllabic: anstatt, aufgrund, ausser, entlang, gegen, hinter, infolge, neben, ohne, über, unter, während, wegen, zwischen
c. polysyllabic: angesichts, ausserhalb, gegenüber, innerhalb

Verbal functional elements (auxiliaries and modals) are mostly monosyllabic in English (69), but forms with the inflectional ending -ing and the contracted negative ending -n't show that disyllabic forms are by no means an anomaly.

(69) Verbal fic in English
a. monosyllabic: be, am, is, are, was, were, been, do, does, did, have, has, had, can, could, may, might, must, ought, shall, should, will, would, can't, won't, don't, aren't, weren't
b. disyllabic: being, having, isn't, wasn't, hasn't, haven't, hadn't, doesn't, didn't, couldn't, mustn't, oughtn't, shouldn't, wouldn't

In German, inflectional endings make many forms disyllabic.\(^35\)

(70) Verbal fic in German
(from the paradigms of sein 'be', haben 'have', werden 'become')

a. monosyllabic: sein, bin, bist, ist, sind, seid, sei, seist, war, warst, wart, hast, hat, habt, wirst, wird
b. disyllabic: waren, wären, wärest, wären, seien, seiet, haben, habe, hatte, hättest, hatten, hattet, hätten, hätten, werden, werde, wertet, wurde, würdest, wurden, würdet, würde, würdest, würden, würdet

The short overview of different-sized function words in (67)-(70) shows that prosodically full-sized forms are not uncommon as functional elements. The difference between foot-sized function words and (monomoraic) syllable-sized function words is that the former but not the latter can be independent prosodic words, as far as the prosodic constraints (HEADEDNESS, FTBIN, PARSE-INTO-f) are concerned. Therefore the direct prosodic argument that the fic cannot be an independent \(\omega\) does not hold for foot-sized cases, and the structure in (71a) is a real contender to (b) and (c) for collocations like under water, haven't eaten, unter Wasser, haben gegessen, etc.

(71) a. fic as independent \(\omega\) b. \(\omega\)-adjoined c. \(\phi\)-attached

Previous research has, however, provided evidence against independent \(\omega\)-hood of foot-sized fic for both English (McCarthy 1993) and German (Kabak and Schiering 2006). In what follows, we will review these arguments against independent \(\omega\)-hood, and show that a broader survey of the same type of facts also turns out to provide key evidence against \(\phi\)-attachment, leaving \(\omega\)-adjunction as the only viable option.

\(^35\) Different from their English counterparts, German modals (dürfen, können, mögen, müssen, sollen, wollen, etc.) act as independent predicates.
4.4.1. English

McCarthy (1993) observes that portmanteau function words such as *shoulda (< should have), gonna (< going to), didja (< did you), lotta (< lot of) do not allow r-insertion (72a), and in this way behave like their monosyllabic counterparts (72b), not like certifiable prosodic words (72c).

(72) a. *gonna-r-eat  b. *to-r-eat  c. Pamela-r-eats
    *lotta-r-apples *
    *the-r-apples California-r-apples

Since these portmanteau function words have the shape of a trochaic foot (stressed σ followed by unstressed ω), it is natural to assume that they are footed and fulfill PARSE-INTO-f. Their footing immediately explains why disyllabic function words do not have reduced variants (cf. monosyllabic have [hæv]~[hæv] vs. disyllabic haven’t [hævənt], *[hævənt] (Selkirk 1984: 355)).

If such foot-sized fuc’s (72a) are independent ω’s as in (72a), r-insertion is incorrectly predicted in either the onset-based approach (section 4.3 above and Ito and Mester to appear) or the coda-based analysis (McCarthy 1993, and Selkirk 1996). In the former, the post-fuc word eat is a maximal ω, hence ONSET-OMAX enforces r-insertion. In the latter, the function word gonna is a prosodic word on its own and receives a final consonant by FINAL-C.

(73) a. independent-ω  b. ω-adjoined  c. φ-attached

We are then left with the familiar choice between ω-adjoined (73b) and φ-attached (73c). The onset-based approach requires the ω-adjoined structure (eat is a maximal ω in (73c) but not in (73b)), whereas the coda-based approach makes no distinction between the two (gonna is not a prosodic word in either (73b) or (73c)).

Previous discussion of r-insertion has not made explicit the exact prosodic status of the function words in such examples. McCarthy (1993) notes that

"[w]e attempt the portmanteau shoulda, gonna, or dija [...] are also proclitic has not been discussed in the literature, but by parity of reasoning they should procliticize too, giving structures like (shoulda eaten)r-wd, (gonna eat)r-wd, and (didja eat)r-wd," and Selkirk (1996) refers to the same types of examples but does not provide specific discussion or structures indicating foot status. 36

A closer inspection of the varieties of English portmanteau function words, however, turns up decisive evidence in favor of ω-adjunction and against φ-attachment. The relevant cases are trisyllabic forms, such as those in (74).

(74) Tri-syllabic portmanteau function words
    supposeta eat  (< supposed to)
    shouldn’t a eaten  (< shouldn’t have, should not have)
    couldn’t a eaten  (< couldn’t have, could not have)
    oughtn’t a eaten  (< oughtn’t have, ought not have)

    These function words, consisting of a disyllabic foot and an unfooted light syllable, must be grouped into some kind of prosodic unit – in fact, the prosodic word, given the ranking FrBIN → PARSE-INTO-f. 37

(75)

Plugging these ω-sized portmanteaus into the familiar three structures, we end up with the configurations in (76).

36 One unresolved issue for either analysis is the mixed behavior of the postverbal reduced pronoun ya (cf. Selkirk 1996: 159–160, notes 14 and 15). In φ-final position, it behaves as a regular object pronoun (see note 33 above). Forming a single prosodic word together with its host verb, it can therefore appear in the reduced form in this position (see ya, goya). On the other hand, in non-φ-final position, ya seems to act as a proclitic to what follows, like other fuc’s, and does not allow r-insertion (give ya up, give ya all, give ya everything, give ya it, saw ya on TV, get ya Amanda.)

37 Recursive foot structure is in principle another option, but would require, besides a new category "superfoot" and concomitant FrBIN violations, the admission of both amphibrachs and dactyls.
Two conclusions emerge. First, the failure of r-insertion after trisyllabic portmanteaus is highly problematic for the coda-based approach (relying on Final-C) since it does not predict the blocking of r-insertion in this context, whichever of the three structures is adopted. The deeper reason for the problem lies in a view of prosodic structure that does not sufficiently distinguish structural subordination and structural reduction. The two often correlate, as in monosyllabic and disyllabic function words, which are non-\( \omega \)'s subordinate to another \( \omega \) (77a,b). But they diverge in longer function words, which are prosodically subordinate (non-heads within a larger maximal \( \omega \)), but still \( \omega \)'s in themselves, as expressed in the right-headed adjunction structure (77c), where the head is underlined. The correct generalization about r-insertion lies in prosodic subordination, not reduction: It is blocked not only after non-\( \omega \)-sized function words, but also after \( \omega \)-sized function words that are prosodically subordinate. This is correctly captured in the onset-based approach, whose pivotal notion is the onset of the maximal prosodic word.

Secondly, we now have evidence against \( \varphi \)-attachment of proclitic function words in English: For trisyllabic function words, the \( \varphi \)-attached structure (76c) turns out to have merged with the independent-\( \omega \) structure (76a), and consequently \( \varphi \)-attachment wrongly predicts r-insertion both for the coda-approach and for the onset-approach. For the latter, advocated here, the \( \omega \)-adjointed structure (76b) makes the correct prediction: Since \( \text{eat} \) is not a maximal \( \omega \), there is no r-insertion. Whether the adjoined prosodic unit is a syllable (77a), a foot (77b), or a prosodic word (77c), the prosodic status of the lexical word \( \text{eat} \) remains non-maximal.

4.4.2. German

Contractions of function word complexes, such as \( \text{auf dem Kanal} > \text{auf'm Kanal} \) (on the canal) or \( \text{wenn es geht} > \text{wenn's geht} \) (if it's OK), are a productive feature of many varieties of contemporary German (see e.g. Duden 2005). They are optional and show significant variation depending on register, degree of formality, and speech rate. They differ in this respect from the small number of obligatory preposition + article portmanteaus such as \( \text{am, ans, im, ins, von, zum, zur} \) (from the prepositions \( \text{an, in, von, zu 'on, in, from, to}' \) and the definite articles \( \text{dem, das, der} \)), which are obligatory in the sense that they block the non-contracted versions at any level of speech (i.e., unless a demonstrative meaning is intended). For example, \( \text{zum Mond} \) 'to the moon' blocks *zu dem Mond, but colloquial \( \text{an 'n Mond} \) (id.) exists alongside \( \text{an den Mond} \). The productive optional contractions are not limited to the prototypical preposition + definite article constellation (78a), but occur with all kinds of \( \text{fnc-fnc} \) sequences, including conjunction + definite article (78b), auxiliary + indefinite article (78c), \( \text{wz-word + pronoun} \) (78d), or conjunction + pronoun (78e).

(78) German \( \text{fnc-fnc} \) contractions

a. \( \text{mit'm [mi.pn.] Rad} \) mit dem Rad 'with the bike'
   b. \( \text{weil's Wetter} \) weil das Wetter 'because the weather'
   c. ich will'n Buch lesen ich \( \text{will'\text{e}} \) Buch lesen 'I want to read a book'
   d. wie's geht \( \text{wie es geht} \) 'how it goes'
   e. wennze [\text{ven.az.}] will wenn sie will 'if she wants'

In a detailed study of the prosodic form of such \( \text{fnc-fnc} \) complexes in several German dialects, Kabak and Schiering (2006) (henceforth K&S) make several interesting points, summarized in (79) and briefly reviewed below.\(^{38}\)

(79) German 

i. The two function words contract to form a trochaic foot [\( f \text{fnc fnc} \)];
   ii. this foot provides the context for specific phonological processes and allomorphs in several dialects;
   iii. the foot-sized \( \text{fnc} \) complexes cannot be independent prosodic words;
   iv. the \( \text{fnc-fnc} \) foot is directly attached to \( \varphi \).

\(^{38}\) Our discussion here is limited to questions of synchronic prosodic form; see K&S (2006) on its historical genesis as well as additional morphological developments (such as reanalysis and analogical extension) that have arguably led to the reinterpretation of the contracted forms as inflected function words.
While we are in agreement with (i)-(iii), we will argue that a closer look at the full range of contracted forms shows that $\omega$-adjunction is preferable to $\varphi$-attachment.

The \textit{fnc-fnc} complexes in question have the surface prosodic shape of a trochaic foot, either as a stressed heavy syllable (80a) or as sequence of stressed-unstressed syllable (80b) where "the initial syllable [...] has more prominent stress and more substance than the subsequent Fnc in the same complex" (K&S 2006: 79).\footnote{K&S (2006) also briefly consider $\omega$-adjunction as an option, but adopt Hall's (1999) position (for the monosyllabic function words) that they are directly attached to $\varphi$.}

\begin{enumerate}[a.]

\item[fnc-fnc feet]

\begin{enumerate}[a.]

\item Consonant deletion in \textit{[fnc-fnc]}:\footnote{Glosses and corresponding full forms: \textit{wenn es} [\textit{ven es}] 'when it', \textit{wenn du} [\textit{ven u:}] 'when you', \textit{wenn er} [\textit{ven er}] 'when he', \textit{für den} [\textit{fyr en}] 'for the (m.acc.)', \textit{weil ein} [\textit{vail ein}] because a (m./n.), \textit{weil eine} [\textit{vail ana}] 'because a (f)', \textit{vor dem} [\textit{vo dem}] 'before the', \textit{wenn sie} [\textit{ven zir}] 'when she', \textit{weil einen} [\textit{vail ama:n}] 'because a (m.acc.).'}

\begin{align*}
\text{\textit{auf'm}} & > \text{'auf'm} \quad \text{`on (m./n.)'} \\
\text{\textit{nach'm}} & > \text{`nach'm} \quad \text{`to (m./n.)'} \\
\text{\textit{mit'm}} & > \text{`mit'm} \quad \text{`with (m./n.)'}
\end{align*}

\item No consonant deletion in \textit{[\textit{f} lex]}:

\begin{align*}
\text{\textit{kaufm}} & > \text{`kaufm} \quad \text{`buy'} \\
\text{\textit{flachem}} & > \text{`flachem} \quad \text{`flat (m.dat.)'} \\
\text{\textit{wippen}} & > \text{`wippen} \quad \text{`swing'}
\end{align*}

\end{enumerate}

\end{enumerate}

K&S discuss several phonological processes (flapping, consonant deletion, consonant intrusion) that occur only in the context of such \textit{fnc-fnc} feet in various dialects. In Ruhrdeutsch (K&S 2006: 75), for example, disyllabic feet can become monosyllabic through medial consonant deletion in [\textit{fnc-fnc}] (81a), but not in [\textit{f} lex] (81b).

\begin{enumerate}[a.]

\item Consonant deletion in [\textit{fnc-fnc}]:\footnote{These cases might be considered as listed allomorphs, but K&S (2006: 68) are careful to point to the existence of a complete sequence of forms at each step of reduction: [\textit{mrt dem}] $\rightarrow$ [\textit{madem}] $\rightarrow$ [\textit{marpm}] $\rightarrow$ [\textit{nmm}] \textit{mit dem} 'with the (m./n.).', "showing the gradual processes of phonological reduction and assimilations'. This makes a strict allomorphy approach less attractive: Not only would all possibilities have to be listed, there would also be no explanation why every level of reduction happens to be lexically precompiled and encoded (what is included in an allomorph list, and what not, is inherently arbitrary).}

\begin{align*}
\text{\textit{auf'm}} & > \text{`auf'm} \quad \text{`on (m./n.)'} \\
\text{\textit{nach'm}} & > \text{`nach'm} \quad \text{`to (m./n.)'} \\
\text{\textit{mit'm}} & > \text{`mit'm} \quad \text{`with (m./n.)'}
\end{align*}

\item No consonant deletion in [\textit{f} lex]:

\begin{align*}
\text{\textit{kaufm}} & > \text{`kaufm} \quad \text{`buy'} \\
\text{\textit{flachem}} & > \text{`flachem} \quad \text{`flat (m.dat.)'} \\
\text{\textit{wippen}} & > \text{`wippen} \quad \text{`swing'}
\end{align*}

\end{enumerate}

K&S's (2006: 75) basic point is that in order to single out [\textit{fnc-fnc}] as a specific reduction site, it cannot constitute an independent prosodic word (82a) since it would then be prosodically indistinguishable from [\textit{f} lex], where reduction is banned. Its special status for reduction (81), as well as for several other \textit{fnc-fnc}-specific processes, such as flapping and consonant insertion, is clearly represented by a prosodic subordination of the foot, whether $\omega$-adjointed (82b) or $\varphi$-attached (82c).\footnote{An alternative approach, which we will leave unexplored here, ascribes the resistance of content words to reduction processes that affect function word complexes not to a specific kind of prosodic representation singling out the latter, but rather to higher-ranking faithfulness constraints associated with the former.}

\begin{enumerate}[a.]

\item独立

\item $\omega$-adjointed foot

\item $\varphi$-attached foot

\end{enumerate}
The extended prosodic word 185

Crucially, there is no difference between disyllabic and trisyllabic fnc-fnc complexes, undermining an exclusively foot-based analysis.

(83b)

Middle Frankish (K&S 2006: 69-70)

a. $\sigma : fnc-r-fnc$
   
   |\begin{tabular}{l}
   [\textit{VOU-R-I}] \\
   [\textit{VOU-R-Y}] \\
   [\textit{gegen-R-an}] \\
   [\textit{zu-einer-R-S}] \\
   [\textit{neben-R-ES}] \\
   \end{tabular} |
   
   |\begin{tabular}{l}
   \textit{wo ich bin} \quad \text{‘where I am’} \\
   \textit{wo er arbeitet} \quad \text{‘where he works’} \\
   \textit{gegen den Klaus} \quad \text{‘against Klaus’} \\
   \textit{zu einer Schule} \quad \text{‘to a school’} \\
   \textit{neben das Haus} \quad \text{‘beside the house’} \\
   \end{tabular}
   
   b. $\ast\text{internal-r-}$
   
   |\begin{tabular}{l}
   [\textit{kar-R-OS}] \\
   [\textit{auder-R-m[en]o\varnothing}] \\
   [\textit{die-Orga\varnothing}] \\
   [\textit{schae ich}] \\
   \end{tabular} |
   
   |\begin{tabular}{l}
   \textit{Chaos} \quad \text{‘chaos’} \\
   \textit{Autoingenieur} \quad \text{‘car engineer’} \\
   \textit{die Orangen} \quad \text{‘the oranges’} \\
   \textit{I watch}
   \end{tabular}

K&S argue that we are not dealing here with an odd, idiosyncratic property of specific items in one specific dialect. Crediting Ortmann (1998) for facts and generalizations, they observe that the same kind of distribution holds for intrusive n ("Binde-n") in Swiss German dialects ("Higher Alemannic"), as seen in (84). Especially noteworthy is the sharp contrast between the last examples in (84a,b), with intrusion before unstressed pronouns (zu-n-ere), but not before emphatically stressed pronouns (zu-*n-IRE). Again, and significantly for our argument here, trisyllabic fnc-fnc complexes pattern with disyllabic ones.

(84) Higher Alemannic (K&S 2006: 72-73)

a. $\sigma : fnc-n-fnc$
   
   |\begin{tabular}{l}
   \textit{woer} \\
   \textit{bei zu-n-ere} \\
   \end{tabular} |
   
   |\begin{tabular}{l}
   \textit{grossier als ich} \quad \text{‘taller than I’} \\
   \textit{als er gekommen ist} \quad \text{‘when he comes’} \\
   \textit{bei ihnen} \quad \text{‘with them’} \\
   \textit{geh zu ihr} \quad \text{‘go to her’} \\
   \end{tabular}
   
   b. $\ast\text{lex-n-fnc}$
   
   |\begin{tabular}{l}
   \textit{China-*n-un Japan} \\
   \textit{wo-*n-igendeine} \\
   \textit{gang zu-*n-IRE} \\
   \end{tabular} |
   
   |\begin{tabular}{l}
   \textit{China & Japan} \quad \text{‘China & Japan’} \\
   \textit{als irgendeine} \quad \text{‘when someone comes’} \\
   \textit{gekommen ist} \quad \text{‘comes’} \\
   \textit{geh zu IHR} \quad \text{‘go to HER’} \\
   \end{tabular}

The argument for $\omega$-adjunction given for English in section 4.4.1 applies here with the same force: Given trochaic foot structure, and $\text{FtBin} \rightarrow \text{PARSE-INTO-f}$, the trisyllabic fnc-fnc must be parsed out as an $\omega$. Only $\omega$-adjunction, not $q$-attachment, can distinguish the fnc-fnc complex in the appropriate way.

4.5. Further consequences

We have seen so far that fnc-fnc contractions in English (\textit{lotta}, etc.) and German (\textit{mit'm}, etc.) are best understood as $\omega$-adjointed. What, then, is the prosodic structure of non-contracted sequences of function words, i.e., those that are not phonologically merged into foot-sized (or in some cases, $\omega$-sized) units? Here, there are two possibilities: one in which the $\omega$-joins each fnc separately (86a), and another in which the two fnc's are parsed into a foot before $\omega$-adjoining them together as a foot (86b). Although further detailed investigation is necessary, our observation is that English in general prefers the former, and German the latter.

(86) a. English preferred fnc-fnc
   
   \begin{tabular}{c}
   fnc-fnc \\
   \end{tabular}
   
   |\begin{tabular}{l}
   *th1 climate \\
   \end{tabular} |
   
   b. German preferred fnc-fnc
   
   \begin{tabular}{c}
   fnc-fnc \\
   \end{tabular}
   
   |\begin{tabular}{l}
   *forth climate \\
   \end{tabular}

The difference in preferred structure captures the following facts: Portmanteaus are more common in German (\textit{für das} $\rightarrow$ \textit{fürs}) than in English (\textit{for the} $\rightarrow$ \textit{forth'}), because fusing the two fnc's segmentally does not disrupt the already existing foot structure. In German, even in non-emphatic/non-focused contexts,
the first fnc, as the initial member of a foot, has prominence (für das ['fyː.das'] ...), whereas strings of stressless fnc-syllables are found in English (für thé ['fr.ðə'] ...).43

Besides the doubly w-adjoined structure in (86a), we briefly consider alternative structures for the unfooted string of syllables, as systematically diagrammed in (87) and (88). In the first three structures (87a–c), two adjacent syllables are attached as sisters to the same higher prosodic node.

(87) Alternate structures I

a. \[ \text{for thé house} \]

b. \[ \text{for thé house} \]

c. \[ \text{for thé house} \]

The next group of structures (88a–c) shows stepwise adjunctions: Two at the w-level in (88a) (=86a), the structure argued for here), one each at the w-level and the p-level in (88b), and two at the p-level in (88c).44

(88) Alternative structures II

a. \[ \text{for thé house} \]

b. \[ \text{for thé house} \]

c. \[ \text{for thé house} \]

Among these, (87a–c) are ruled out as gratuitous violations of PARSE-INTO-f,45 and although (88a–c) are all viable options for representing a sequence of two unstressed syllables,46 the intrusive-r- evidence turns out to favor (88a), as shown in (89) with examples of hiatus sites in different positions.

(89) maximal: \[ \text{for thé house} \]

b. \[ \text{for thé house} \]

c. \[ \text{for thé house} \]

Only the doubly w-adjoined structure (89a) declares the entire fnc-fnc-lex complex to be a maximal w (circled), and correctly predicts r-insertion in law-[r-ôf thé land], but not in [tô.(h)is troubles] nor in [for thé apples], where the hiatus sites lie before nonmaximal w’s. On the other hand, the mixed adjunction structure (89b) declares the inner fnc-lex complex to be a maximal w, wrongly predicting r-insertion in tô-[*r-is troubles]. The p-adjunction structure (89c) does the same for the innermost lex, yielding the incorrect for thé-[*r-apples].

Thus, both rhythmic form (unstressed sequence of syllables) and segmental form (r-intrusion) point to the doubly w-adjoined (90a), which is left to compete with (90b) where the two fnc’s are first parsed into a single foot and w-adjoined together.

(90) a. \[ \text{fnc}_1 \]

b. \[ \text{fnc}_1 \]

c. \[ \text{fnc}_1 \]

Both structures fulfill PARSE-INTO-w, but (90a) (with fnc2 parsed twice, and lex parsed three times) violates NO-RECURSION to a greater degree than (90b) (lex is parsed only twice). So, all else being equal, (90b) should always be preferred. What forces the doubly adjoined structure, we suggest, is a HEAD-TO-LEX constraint (91) requiring heads of prosodic categories to be contained in lexical (not functional) material.

46 As discussed in section 3.1 (see (35)) proper bracketing blocks cross-boundary footings in (88a–c).

44 Both languages have trochaic feet. Note that there is no systematic vowel reduction in German, so unstressed syllables retain their vowel quality.

45 (88c) violates NO-RECURSION at the p-level. The mixed adjunction structure (88b) is an interesting alternative to consider, since it might even suggest that each level is limited to a single adjunction. But no such restriction seems to be empirically valid (cf. and for thé house, etc.).

46 But see Peperkamp (1996) who makes a distinction between lexical and postlexical foot parsing.
(91) HEAD-TO-LEX: Prosodic heads must be contained in lex.

HEAD-TO-LEX is fulfilled in (90a), where both fnc's avoid headhood by adjunction, and remain weak unstressed syllables.47 On the other hand, it is violated in (90b), where the head of the adjoined fnc-fnc foot lies outside of lex (the violation would be especially notable if the fnc-fnc foot constituted a subordinate prosodic word, as in the trisyllabic cases (77c)). The choice of optimal structure will then depend on the ranking: HEAD-TO-LEX >> NO-RECURSION will yield (90a), while the reverse ranking will make (90b) optimal. The HEAD-TO-LEX constraint straightforwardly captures the preferred non-prominence of function words, and might be considered part of the rationale behind their general invisibility.

5. Conclusion

We have presented some thoughts towards a new conception of the constraints concerning layering in prosodic structure. Much is to be gained, in our view, by adopting a slimmed down version of the prosodic hierarchy beyond the word level, with fewer categories but with a richer projection structure where adjunction plays an important role. We have put these ideas to work in considering some questions regarding the phonology of clitics. At least for the languages closely considered in this work, English and German, there is a large amount of evidence that the prosody characterizing a typical function word is not as impoverished an item that has not attained prosodic word-hood and is directly attached to the phonological phrase, but rather as a dependent element within an extended prosodic word structure, where it has adjoined to a core prosodic word, its head.

Acknowledgements. Portions of this work have been presented at the 4th Old World Conference in Phonology (Rhodes, Greece, Jan. 2007), the 2nd Atami Phonology Festa (Atami, Japan, Feb. 2007), the 3rd workshop on Prosody, Syntax, and Information Structure (Indiana University, Sept. 2007), in colloquia at UC Santa Cruz (Oct. 2007) and at UCLA (Jan. 2008), and in class lectures at Kyoto University (Spring 2007) and at UC Santa Cruz (Fall 2008). For discussion and helpful comments, we are grateful to the audiences at these occasions, in particular to Outil Bat-EI, Mary Beckman, Ricardo

47 A straightforward alignment constraint aligning left edges of fnc to ə also produces the doubly-o-adjoined structures, but goes against the generally accepted view that syntact-phono mapping constraints do not refer to functional categories (Selkirk's 1984: 343 `Principle of the Categorial Invisibility of Function Words`). Alternatively, one might appeal to some requirement that prosodic structure should closely mirror the syntactic adjunction structure (cf. van Oostendorp's 2002 MIRRORING principle).

References


Flack, Kathryn 2006 Restrictions on the edges of prosodic domains. Ms. Amherst: University of Massachusetts, Amherst.
Fougeron, Cécile and Patricia Keating

Gelfiu-Wolfgang, Jochen

Giegerich, Heinz

Gutch, Donald

Gouskova, Maria

Goldsmith, John

Gutek, Donald

Hall, Tracy Alan

Hall, Tracy Alan, and Ursula Kleinhenz (eds.)

Halliday, Michael A.K.

Harden, Theo

Hayes, Bruce
1984 The prosodic hierarchy in meter. Ms. UCLA.

Hyde, Brett

Ikegami, Sharon

Ito, Junko and Armin Mester


Jurafsky, Daniel, Alan Bell, Michelle Gregory and William D. Raymond

Kabak, Barış and René Schiering

Kahn, Daniel

Kiparsky, Paul

Kleinhenz, Ursula

Kubozono, Haruo
Ladd, D. Robert  
1986  
1996  
*Intonational Phonology*. Cambridge, UK: Cambridge University Press.  

McCarthy, John J  
1991  
1993  

McCarthy, John J, and Alan S. Prince  
1993a  
1993b  
Prosodic morphology I: Constraint interaction and satisfaction. Ms.: University of Massachusetts, Amherst, and Rutgers University.  

McMahon, April M. S.  
2000  

Meinschaefer, Judith  
2005  

Neef, Martin  
1996  

Nespor, Marina  
1990  

Nespor, Marina, and Irene Vogel  
1983  

1986  
*Prosodic Phonology*. Dordrecht: Foris.  

Oostendorp, Marc van  
2002  

Ortmann, Albert  
1998  

Pater, Joe  
2000  

Peperkamp, Sharon  
1996  
1997  

Pierrehumbert, Janet and Mary Beckman  
1988  

Prince, Kenneth  
1967  

1980  

1983  

Prince, Alan and Paul Smolensky  
2004  

Selkirk, Elisabeth  
1981  

1984  

1986  
1996  

2000  

Selkirk, Elisabeth and T. Shen  
1990  
Interface Explorations 16

Editors
Artemis Alexiadou
T. Alan Hall

Phonological Domains
Universals and Deviations

edited by
Janet Grijzenhout
Barış Kabak

Mouton de Gruyter
Berlin • New York
Contents

List of Contributors ......................................... vii
Acknowledgements ............................................... ix

Prosodic Phonology: An appraisal .......................... 1
Janet Grijzenhout and Barış Kabak

Part 1. Prosodic Hierarchy and the Nature of Prosodic Constituents

The status of the Clitic Group ............................... 15
Irene Vogel
The distribution of phonological word domains: A probabilistic typology 47
Balthasar Bickel, Kristine A. Hildebrandt, and René Schiering

Part 2. Adjunction, Recursion, and the Nature of Syntax-Phonology Mapping

Phrase-level and word-level syllables: Resyllabification and prosodization of clitics 79
Anna Cardinaletti and Lori Repetti
An interface approach to prosodic word recursion 105
Barış Kabak and Anthi Revithiadou
The extended prosodic word ................................. 135
Junko Ito and Armin Mester
Multiple spell-out, assembly problem, and syntax-phonology mapping 195
Yoshihito Dobashi
The role of Weight-by-Position in the prosodic development of Spanish and German 221
Conxita Lleó and Javier Arias
Decomposition of question intonation: The structure of response seeking utterances 249
Aslı Göksel, Meltem Kelepir, and Aslı Üntak-Tarhan