Spans and Words

Peter Svenonius
*CASTL, University of Tromsø – The Arctic University of Norway**

Published in 2016 in *Morphological Metatheory*, ed. by Heidi Harley and Daniel Siddiqi, 199–220. John Benjamins, Amsterdam.

Abstract

1. The problem: Words are a pervasive unit of syntax and yet the dominant theory of them, the X^0 theory, is problematic, predicting more parallels between phrasal and head movement than are observed. Phrasal movement approaches to word formation fare even worse on that score. Mirror Theory (MT) also has shortcomings, for example in relying on an unmotivated notion of specifier.
2. The solution: A theory of how syntactic structures are mapped onto functional and lexical words, positing syntactic features w for lexical access points and @ for linearization points. The theory draws on the late insertion of DM, the cycles of phase theory, the direct linearization principles of MT, and the non-terminal spell-out of Nanosyntax, separating word formation from linearization and appealing to *spans* (head-complement sequences) as the units of cyclic lexical access and storage.

1 Introduction: In search of the grammatical word

Phonological analysis more or less universally recognizes a category Phonological Word (sometimes abbreviated PhWd or ω) as a domain for certain

*Thanks to Ricardo Bermúdez-Otero and Terje Lohndal for valuable discussion and to two anonymous reviewers for insightful and helpful comments.

processes and phenomena, for example stress placement, word-final devoicing, vowel harmony, minimal word constraints, and so on.[1] In fact, the word appears to be the maximal domain for cyclic phonology other than the prosodic structure produced by syntax; as Bermúdez-Otero (2011) puts it, "there appear to be no cyclic nodes between the grammatical word (X^0) and the utterance"; hence the phonology applying across words is a different process from the phonology applying within words (however, in contrast to the presupposition in the quote, I will argue in this paper that the grammatical word cannot be equated with X^0).

Given a string of segments, purely phonological considerations can be applied to determine syllable structure and foot structure. But no purely phonological considerations can be applied to determine word boundaries. These are externally imposed.[2]

Thus, phonology relies on syntax to identify the word. Therefore it is a problem for syntax if syntacticians rely in practice on phonological characteristics to identify a grammatical word.

In a lexicalist model, the word is the output of the morphological component. Morphology, almost by definition, recognizes a distinction between word-internal and word-external phenomena; allomorph selection across a word boundary is limited to syntactic and phonological features, while allomorph selection within a word can also be sensitive to class and lexical features.

Thus, the word is central to lexicalist models, but those models have not developed any theory explaining why words have the boundaries that they do.

In a late-insertion model like DM, in which there is no distinct computational word formation module, the problem is just as acute. A theory of the grammatical word is needed that can feed the phonology. Until then, the word is at best a cryptophonological diacritic in the syntax telling the syntax how to spell something out, with no account for the generalizations which

[1] Though Scheer (2008) argues that only word *boundaries* are motivated, not *domains*.

[2] Déchaine (2015) discusses the interesting case of Yorùbá, where verbs are restricted to CV syllables, and other lexical words are longer; thus, descriptively, one could say that a lexical CV word is a verb. Alternatively, there is a CV template associated with the category verb (perhaps it is the phonological structure of a category-defining head v). What I am pointing out here is that there is no language in which strings of segments are broken up into words on the basis of purely phonological considerations, the way syllables and feet can be constructed on purely phonological grounds.

are observed (Scheer 2008). I discuss the X^0 hypothesis of the grammatical word favored in DM in §3.

It is sometimes thought that word boundaries are essentially arbitrary. For example, *smarter* has no word boundary between the comparative morpheme and the root, while the functionally equivalent *more intelligent* does,[3] and the English (1a) contains three lexical words and two function words while the semantically similar Nahuatl example in (1b) is a single word.

(1) a. I want to eat potatoes.
b. Ni-kamo-kwa-s-neki.
1SG-*potato*-*eat*-SUB-*want*
'I want to eat potatoes' (Tenextatiloyan Nahuatl, from own 1992 field notes)

But despite some variation, the boundaries are not arbitrary. For example, there are crosslinguistically valid constraints on the interpretation of incorporated versus free nouns, consistent with the claim of Baker (1988) that what is incorporated is a projection of N, not of D. For example, an incorporated noun like *kamo* 'potato' in the Nahuatl example can be characterized as lacking a D projection, and hence typically interpreted as a weak, narrow scope indefinite (Carlson 2006). Nahuatl cannot incorporate the DP *in kamo* 'the potato.'

Furthermore, finite verbs do not incorporate, only non-finite ones like *kwa* 'eat' in (1b). I have glossed the *-s-* there as SUB[ordinate]; its exact category is unclear. But it is not a finite ending, and a full finite form could not be incorporated into the higher verb *neki* 'want.'

It seems that higher-level functional categories such as DP and CP never incorporate into the lexical material under which they are subordinated; this in itself shows that word boundaries are not entirely arbitrary. The universal generalization would be that C and D induce word boundaries; the parametric property would be which of the other categories also do.

In that case we can say that unincorporated phrases have additional material, compared to incorporated phrases (compare Cardinaletti and Starke's 1999 arguments that clitics are structurally deficient in a way that restricts

[3]Though, as is well known, analytic phrases with *more* have additional uses not shared by the synthetic forms, e.g. the 'metalinguistic' use in *It is* {*more cold/*colder*} *than hot* (Di Sciullo & Williams 1987). This situation is not uncommon in cases of periphrasis, see for example Nikolaeva (2013) for an example from Tundra Nenets.

their interpretation as well as being reflected in their reduced prosody).

The fact that Nahuatl *neki* 'want' incorporates lower verbs can be treated as a matter of subcategorization: *neki* selects a complement which is small enough to incorporate (as a matter of fact it can also take larger complements, but this kind of optionality is common for selection; compare *believe* allowing both ECM infinitives and finite complements in English).

Another source of variation comes from the phonology. Syntax is mapped onto prosodic structures which are the input to phonological computation. The phonology of different languages can treat the same input in different ways. For example, the mapping from syntax to phonology might produce an adjoined prosodic word structure, where one prosodic word is embedded inside another. This violates a general markedness constraint against adjoined structures, in OT terms (Selkirk 1996, Ito & Mester 2012). If faithfulness to the input outranks the constraint against adjunction, then a phonological word boundary will be present in the output.

In English, phonological words may be recursive when lexical words are adjoined to each other, suggesting that in such contexts, faithfulness to the syntactically-determined input outranks the markedness bias against recursive prosodic categories. Compare the phrasal *philosóphical dissertátion*, with phrasal stress on the head noun to the right, with the compound *philósophy dissertàtion*, in which the primary word stress falls on the leftmost member; nevertheless, the underlying stress patterns of both words are preserved.

(2)

								x	
			x					x	
	x		x			x		x	
phi	lo	so	phi	cal	dis	ser	ta	tion	

(3)

		x							
		x					x		
		x			x		x		
phi		lo	so	phy	dis	ser	ta	tion	

If, on the other hand, the constraint against adjunction outranks faithfulness to the input, then the internal word boundary might be erased. This is arguably the case in Central Yupik, where noun incorporation does not show a word-internal word boundary on the surface. Woodbury (1987) gives the minimal pair in (4), where all but the first morpheme are identical. Iambic feet are constructed over the entire word from the left, which causes a stress

to fall on the verb root in (4a) (and every other syllable), but not in (4b), where the incorporee is disyllabic.

(4) a. pi-ssú:-tu-llí:ni-lú:ni
thing-hunt-always-apparently-APP.3REFL
'S/he apparently always hunted [things]'
b. mallú:-ssu-tú:-lliní:-luni
beached.whale-hunt-always-apparently-APP.3REFL
'S/he apparently always hunted beached whales'

In this paper, I propose a novel theory of the word, building on phase-based DM as well as Brody's (2000a, 2000b) Mirror Theory, which is based on the notion of *span*.[4] In short, minimal words are created in the input to phonology when lexical access targets certain nodes for the spell-out of the spans they dominate, where the span is the sequence of complements (going down the extended projection, ignoring specifiers and adjuncts). The notion of span plays another important role in the theory, as exponents are inserted into spans, rather than into terminal nodes. Thus morphemes are exponents of spans, and syntactic words are combinations of morphemes in certain configurations also defined in terms of spans. This is novel because the notion of span has no status in current theory. Using spans eliminates the need for head-movement, fusion, local dislocation, and several other mechanisms commonly used in DM.[5]

2 Proposal: Words are defined over spans

I will assume a late insertion model like the one argued for in Distributed Morphology (DM). However, DM is generally committed to taking syntactic terminal nodes to be the targets of vocabulary insertion — hence a syntactic head canonically corresponds to an exponent, allowing DM to use the term "morpheme" for syntactic heads. I propose instead that vocabulary inser-

[4]Additional related work includes Ramchand (2008), the papers in Svenonius et al. (2009), Adger et al. (2009), Bye & Svenonius (2012), Svenonius (2012; 2016), and Adger (2013).

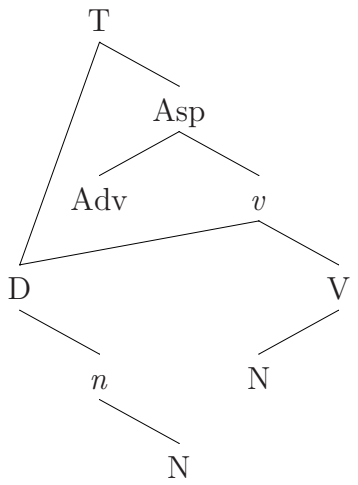
[5]Compared to Nanosyntax (e.g. as represented in Svenonius et al. 2009), spelling out spans eliminates the need for phrasal movements to move specifiers out of targets for insertion, and subsequent movements to relinearize those targets in order to undo the unwanted effects of the specifier evacuation.

tion targets *spans*, which are head-complement sequences, normally in an extended projection.[6]

(5) A span is a contiguous sequence of heads in a head-complement relation

For example, take the following tree, intended to represent a TP clause where the DP subject has moved from v to T, there is an adverb adjoined to AspP, and the object is an NP which has remained in the VP.[7] This might represent a sentence like *The tigers always take note*. Following a convention initiated by Brody (2000a;b), specifiers slope down to the left while complements slope down to the right, so the N dependent of V is a specifier, not a complement, in this tree.

(6)



This tree contains four extended projections: Adv, the subject DP, the object NP, and the clause itself. <Adv> and <N> are, in addition to being extended projections, also trivial spans, as are all other singleton categories. <T, Asp, v, V> constitute a span, and so do all contiguous subsequences: <T, Asp>, <Asp, v>, etc. The DP similarly forms various spans, all subparts of <D, n, N>.

[6]The status of spans which cross extended projections falls outside the scope of this paper. See Son & Svenonius (2008) for a proposal in which P complements of V form a lexicalizable span with the verbal material.

[7]This is a tree in the linguist's sense, not the mathematician's, for whom it is still a graph but not a tree due to the multidominance representation of movement.

Phase theory (Chomsky 2001; 2008, inter alia) holds that certain designated nodes, called phase heads, trigger the operation of Transfer, in which selected phrases are transferred to interface modules for phonological and semantic interpretation. Assuming Late Insertion, part of this Transfer process involves lexical insertion. But a phase may contain specifiers and adjuncts which are distinct targets for lexical insertion.

In the two-stage lexical insertion model of Bye & Svenonius (2012), lexical insertion is divided into a syntactic step of Lexical Matching, where lexical items are matched to syntactic representations according to syntactic features only, and a phonological stage of Insertion, where phonological considerations determine allomorph selection. We propose that the step of Lexical Matching targets spans, rather than terminal nodes or heads.

(7) Lexical matching applies to spans

This means that, within each phase, the spell-out procedure must process each extended projection separately. If the VP is a spell-out domain in (6), then the trivial span V and the trivial span N will be targeted separately for lexical insertion. If the TP is a spell-out domain, then the clausal spine and the subject will be targeted separately. The fact that there is no syntactic constituent T-Asp-v-V which excludes the subject (and the adverb, and the object) is no hindrance to lexical insertion, I contend.

In this section I discuss some of the consequences of (7), and in §3 I compare it to the X^0 hypothesis of word formation.

2.1 Portmanteaux

Examples like **John is an orphan and he misses them* (Postal 1969, * on coreference between *them* and the implicit 'parents') have been used to argue that semantic subparts of morphemes are inaccessible. Syntactically visible, morphemic subparts of complex words, are also often inaccessible, as illustrated by DiSciullo and Williams' (1987) *I babysat last night. #Boy was she ugly.*

However, the inaccessibility of subparts of words is not total. Baker (1996) shows that incorporated nouns in Mohawk can have relatively high degrees of referentiality, compared to subparts of English compounds, and especially compared to single morphemes, crosslinguistically.

Assuming that referentiality correlates with structure, the restriction on

words like *orphan* shows that there are structures which cannot be contained within morphemes. Since any individual functional head can presumably be expressed by a morpheme, this means that it is a structural configuration that cannot be so expressed. In fact, morphemes are so limited in their range of meanings that the standard assumption is that they spell out only X^0, syntactic atoms. But portmanteau phenomena abound, especially if cartography is on the right track, and lexical decomposition suggests that many words have a richer internal structure.[8]

Nevertheless, the impossibility of referring expressions being contained inside morphemes appears to be absolute. Partly, this might follow if morphemes cannot include material from two phases (I suggested this for idioms in Svenonius 2005, building on Marantz 2001). Restrictions on morpheme meaning seem to be stricter still, and so I suggest that they are restricted to spans. The reference to 'single' here is to distinguish these from multimorphemic idioms, which may be inserted into multiple spans.

(8) Single listed exponents are inserted into spans

Thus, for example, if a span can consist of a nominal root, a category head *n*, and a classifier node Cl (taking Cl to be the locus of plural marking following Borer 2005), then a lexical entry like the one for the plural noun *mice* could include all three heads in its syntactic specification, so that a partial lexical entry might appear as follows.

(9) /maɪs/ ⇔ <N, *n*, Cl>

The exponent /maɪs/ could then be inserted directly into a syntactic structure containing the span N–*n*–Cl. This obviates any need for a Fusion operation as used in DM (cf. Siddiqi 2009). Fusion overgenerates as it would allow, for example, a direct object D to fuse with a selecting V, permitting suppletive V-D for certain verbs, a situation which I believe to be unattested.

Insertion into spans is distinct from the nanosyntactic insertion of lexical

[8]The issue of readjustment rules (Halle & Marantz 1993) is relevant. If there are readjustment rules, then forms like *would* and *could* can be 'adjusted' versions of *will* and *can*, and show nothing about what single morphemes can spell out. However, if there are no readjustment rules, then the lack of any transparent segmentation in *would* and *could* suggests that they are portmanteaux, including the features of *will* and *can* plus, at least, the past tense. Examples like this abound, so readjustment and portmanteaux are competing explanations for a large class of phenomena.

entries into syntactic constituents (e.g. Caha 2009), because syntactic constituency is not identical to spanhood. There are constituents which are not spans (for example any projection containing a specifier) and spans which are not constituents (a span is not a constituent which excludes a specifier, if that specifier is a dependent of a head in the span).

The same span N–*n*–Cl could be lexicalized by multiple exponents, for example *pig-let-s*, given lexical entries like those in (10)

(10) a. /pɪg/ ⇔ <N>
b. /lət/ ⇔ <*n*>
c. /z/ ⇔ <Cl>

In both cases (9) and (10) there is a single word. The surface order of the exponents mirrors the hierarchical structure of the syntax, as discussed in §2.3 below.

2.2 Syntactic Periphrasis

Syntactic periphrasis involves a paradigmatic opposition between one word and two. Classic examples are cases where simple tenses alternate with compound tenses. For example, in Turkish, the definite past is a simple tense, while the reported past is a compound tense (Kornfilt 1997), involving a participle and a cliticized present tense copula (cliticization is indicated in (11b) using an equals sign). Though the copula itself is normally null before a consonant, the compound tense is evident in several ways, including word stress (word-final in Turkish) and in the agreement allomorphy (second person singular is *-n* on main verbs, but is *-sin* with the present-tense copula).

(11) a. git-tí-n
go-DEF.PAST-2SG
'you went (definite past)'
b. git-míş=sin
go-REP.PAST=COP.2SG
'you went (reported past)'

The definite past can occur outside the *-miş* participle, forming a past perfect.

(12) git-míş=ti-n
go-REP.PAST=COP.DEF.PAST-2SG
'you had gone'

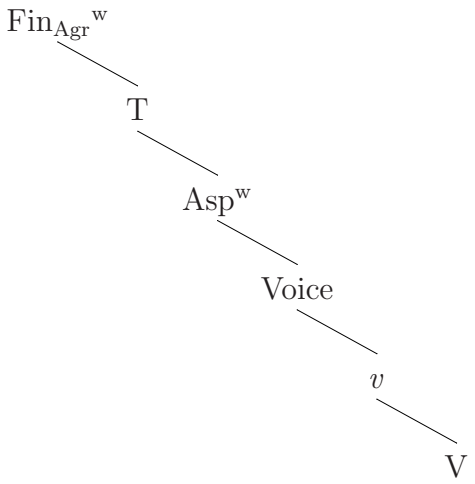
Take *-miş* in the examples in (11b) and (12) to be an exponent of an (outer) Aspect category Asp, which is absent from the definite past form in (11a). Asp bears a property which causes it to spell out as a word distinct from the functional material above it.

The examples in (11) can be represented schematically as in (13), positing one node for each overt morpheme and marking the node inducing wordhood with a superscripted w. The entire word in (11a) is a word, so Fin_{Agr} is marked with w. In (11b), the participle is a prosodic word, but the auxiliary, containing Fin_{Agr} is not a prosodic word. Thus, a surface prosodic word requires something in addition to w, namely a root or stem. Spans marked with w which do not include a root or stem will be morphologically wordlike, but without the extra prosodic prominence of a lexical word.

(13) a. Definite past: V-T-$\text{Fin}_{\text{Agr}}{}^{\text{w}}$
b. Reported past: V-Asp^{w}=T-$\text{Fin}_{\text{Agr}}{}^{\text{w}}$

The string in (13b) corresponds to a syntactic structure like that in (14), positing a few more categories.

(14)

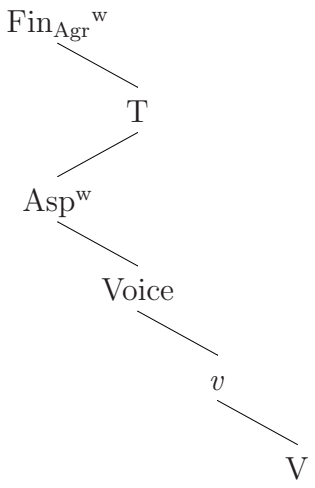


I suggest that a category marked w has a feature which causes it to be a 'cyclic node' for lexical insertion, in the sense of Embick (2010). Effectively,

periphrasis is forced by the cycle of lexical access which occurs at the w mark.

Brody (2000a) suggests that all word boundaries indicate specifierhood, that is, all words are specifiers. In that case, he is forced to stipulate that a participle like that in (11b) occupies a specifier position. This would correspond to the following tree structure.

(15)



So far no independent motivation for the specifierhood of such words has emerged (e.g. difficulty of extraction, reconstruction effects, stranding phenomena), so I will simply mark the category corresponding to a word with w, assuming that it can bear that feature and remain structurally in complement position, contra Brody. Nothing appears to hinge on this.

Since the minimal word is also a domain for stored idiosyncracies of form (e.g. irregular plurals) and content (e.g. pluralia tantum, and what Marantz 2013 calls 'allosemes'), I suggest that it is relevant for lexical access. Thus, the phase may determine the cycle for Transfer, but the span within the phase is the domain which is compared directly to the standard lexical entries, i.e. the morphemes.

(16) Lexical matching is bounded by certain nodes with the property w

For example, in some languages there is a plural marker which is not contained within the same phonological word as the noun, as in the examples below from Galela and Tongan.

(17) a. o bi gotta
ART PL *tree*

'trees' (Galela, Rijkhoff 2002:110)

b. ha fanga pulu
INDEF PL *cow*
'some cows' (Tongan, Dryer 1989:875)

Such languages would differ from English in having a w feature below Cl, for example on n (note that I posit no w on Cl in these languages, as Cl is typically not morphologically complex, unlike the Turkish auxiliary, and there is no alternation with a morphologically complex N-n-Cl, again unlike the Turkish alternation of participial verbs with inflected finite verbs).[9]

(18) a. English: N–n–Cl^{w}
b. Tongan: N–n^{w}–Cl

The w feature is in a sense the inverse of a feature triggering head movement. Head movement, as movement, is normally assumed to be triggered. In a Mirror-based theory, heads in a word are morphologically incorporated by default, unless a feature such as w prevents incorporation. I discuss head movement further, briefly, in §3.

2.3 Mirror theory

I have suggested that w marks the boundary of a lexical word, and that it induces lexical access. Thus it follows that that node, and all of the nodes within the span that it takes as a complement, form a lexical word.

(19) A root and the span dominating it up to the first w feature is a lexical word

The phonological consequences of the w feature must follow from the way that syntax is mapped to prosodic structure. The boundary of the lexical word is interpreted by phonology as the boundary of a phonological word (cf. Selkirk 2011). This should not have to be stipulated, but should be made

[9]The extended projections are represented in (18) from bottom to top, in keeping with the diagrams previously provided for Turkish; equivalently they could be represented as follows:

(i) a. English: Cl^{w} > n > N
b. Tongan: Cl > n^{w} > N

to follow from the effect that cyclic lexical access has on the phonological representation (see Scheer 2008 and Šurkalović 2013 for relevant discussion).

(20) The edge of a lexical word is interpreted in the input to phonology as the boundary of a phonological word

As already noted, a lexical word in this sense may consist of several exponents. There is a major crosslinguistic bias toward suffixation as opposed to prefixation, so that an exponent representing higher material in a span, if it is affixal, normally follows exponents of lower material. This is axiomatic in Williams' (1981) Right Hand Head Rule and in Brody's Mirror Theory, and I will take it as axiomatic here as well, though it should ultimately be explained.[10]

Multiple w's as in the Turkish reported past (cf. (13b)) divide an extended projection up into parts; call each span with a w at the top node, all the way down to the next w if there is one, a "w-span." Linearization of morphemes within a w-span follows the following principle.

(21) If A dominates B in a w-span, then an exponent of A follows an exponent of B

The statement in (21) can be taken as a default, which can be overridden by the same marked phonological alignment features which govern infixation (Bye & Svenonius 2012).

On syntactic approaches to word formation, the linearization of morphemes and of phrases follow the same principles (most prominently Kayne's 1994 LCA). In DM, the prime engine driving complex word formation is head movement. Head movement is syntactic movement and like other kinds of movement is exclusively upward. Therefore, there is a tight connection between the formation of complex words and head raising.

The Mirror-Theoretical alternative is that word formation is not generally driven by head movement, but by complementation (taking traditional complements like direct objects to actually be specifiers of low parts of an extended projection). If dependency in syntax is mapped to precedence in phonology, then specifiers will precede heads, and so will complements.

This gives a default head-final order, with bound morphemes appearing

[10]Compare Brody's (2000a:42) *Mirror*: "The syntactic relation "X complement of Y" is identical to an inverse-order morphological relation "X specifier of Y." This use of the term 'specifier' does not seem to have currency outside Mirror Theory.

as suffixes. Languages vary according to which categories, if any, are head-final. Brody (2000b) proposes that a head in a span may be designated—with a feature he represents with @—as the point at which spell-out inserts the word; specifiers of heads higher than @ will precede the word linearly, but specifiers of heads lower than @ will follow it.[11]

(22) The exponence of a span is linearized within the phrase it projects according to the spell-out feature @

Following Brody, the fact that a span V–*v*–T linearizes in French in the T position indicates that T in French bears the linearization feature @; in English, the linearization feature is lower, on *v*.

(23) a. French: V–*v*–T@
b. English: V–*v*@–T

This means that the English verb will precede anything contained in or adjoined to VP, but will follow specifiers and adjuncts to *v*P and all material in T; while the French verb will precede anything contained in or adjoined to *v*P.

Nespor et al. (2008) suggest that learners can use prosody to determine headedness of phrases; on the present model this means that a learner could posit a high position for @ on the basis of stress-final phonological phrases.

Non-bound function words which linearize independently of the lexical head of their extended projection must involve a distinct @. For a higher function word, the default linearization will be to the left, on the assumption that higher @'s are linearized to the left of lower ones. As a learning heuristic, a learner identifying a head-initial free function word in a phrase need only posit an @ on that category.

In a language like Tongan, where the plural marker precedes the noun, there are two points for linearization in the noun phrase below the determiner,

[11] Bye & Svenonius (2012) discuss the phenomenon of antritropal exponents, with lexical linearization specifications, such as the Latin conjunction *que*. Apart from such cases, linearization is not tied to exponents, but to categories; e.g. verbs can be head-initial or head-final according to their morphosyntactic featural status, but not according to lexical features like conjugation class. This means that the feature @ does not appear to be part of the phonology of lexical entries, but rather resides in the syntax. Linearization itself is a property of phonological representations, not syntactic ones, so in a strictly modular syntax–phonology division the distribution in the syntax of features influencing linearization should ultimately be derived from something deeper.

one for the plural word and one for the noun.

(24) a. English: N–*n*–Cl^{w}@
b. Tongan: N–*n*^{w}@–Cl@

In general, a structure with more @'s than w's will contain function words. In general, languages with little morphology (such as Vietnamese, Noyer 1998) are characterized by having many @'s in their extended projections.

A structure may also contain more w's than @'s. This would mean that there are multiple phonological words, but they linearize together, as with compounds and other cases where one phonological word is adjoined to another, as in the Turkish case discussed above. Such cases may be affected by phonological processes, rendering the underlying structure opaque, as discussed for Central Yupik in §1 above).

Higher prosodic categories often entail lower ones, due to phonological principles. If the prosodic interpretation of a phase is a phonological phrase (Kratzer & Selkirk 2007, Adger 2007, Kahnemuyipour 2009), and DPs or KPs are phases, then DPs or KPs will be phonological phrases in the input to phonology, in which case phonological constraints alone may determine that they are normally at least words in the output of phonology, whether or not they bear an @ feature.

It is unclear what semantic content is associated with @; see Roberts (2010) on the evidence for semantic content to head movement, which would translate into the semantic content of @ on this proposal.

2.4 Grammatical words

The definition of lexical word in (19), as a root plus the incorporated functional material in its extended projection, corresponds to the usual intuitions about what a minimal lexical word is. However, the minimal lexical word does not figure prominently in many syntactic phenomena. What is more important to syntax is the grammatical word, which includes elements that are larger than a minimal lexical word. For example, compounds may contain more than one lexical word, and yet may undergo head movement, as illustrated in (25).

(25) Derfor sitte-danser de attmed bordene sine.
*therefore sit-dance they by tables.*DEF REFL.POSS
'Therefore they sit-dance (dance while sitting) at their tables'

Grammatical words may also include clitics which are attached outside the lexical word, as in the French example in (26), indicating the attached clitic with an equals sign.

(26) Je le=regarde tout simplement.
I it=look.at all simply
'I'm just simply looking at it'

In addition, grammatical words may include function words such as auxiliaries which are not lexical words. For example, the copula in English undergoes verb movement as a grammatical word, but is not a lexical word and does not normally show the full prosodic prominence of a lexical word.

(27) What are you looking at?

Grammatical words, in this sense, are units for syntactic linearization. Therefore, they must be defined in terms of the distribution of @. Clitics and adjoined or incorporated lexical words in compounds lack independent @, by hypothesis. Thus each grammatical word contains exactly one node marked by @. A span containing one @, up to the next @, can be called an @-span. An @-span which is a dependent of an @-span (e.g. a specifier with @, somewhere in an @-span) is a distinct @-span, but a dependent containing no @ will be linearly inseparable from the @-span containing it. Thus, a grammatical word can be defined as the maximal structure containing at most one @.

The distribution of w and @ gives four categories of span, as illustrated in the following table (abstracting away from the properties of roots).

(28)

	@	no @
w	freestanding lexical word	incorporated word
no w	freestanding function word	(internal) clitic

An affix, like a clitic, lacks w and @, but is part of a larger span. Selkirk (1996) identifies three prosodic classes of clitic: free (outside the prosodic word), affixal (adjoined to the prosodic word), and internal (contained in the prosodic word). A case like the Turkish auxiliary, with an independent w but no independent @, would be an affixal clitic, in Selkirk's terms.

Roots have an effect on prosody, perhaps by virtue of triggering a cycle of stem-level phonology, but there seems to be some crosslinguistic variation concerning the prosodic interpretation of roots and their interaction with the

w feature. This is an ongoing topic of research.

3 The X⁰ hypothesis

The most widely held theory of word structure is some version of an X^0 hypothesis, namely that words are X^0s, including possibly complex X^0s. The main engine forming complex words is head movement.

(29) Lexical insertion targets terminal nodes, after head movement, lowering, fusion, and local dislocation have applied

(30) Lexical insertion occurs phase-by-phase, starting with the most deeply embedded node

(31) A maximal X^0 containing a lexical root is a grammatical word

(32) A grammatical word is interpreted in phonology as a phonological word

(33) A maximal X^0 which does not contain a root is a function word

(34) The exponence of an X^0 is linearized according to the LCA

Since head movement is constrained by the Head Movement Constraint (Travis 1984, Baker 1988), the complex heads which it creates are restricted to contiguous heads in a head-complement sequence—precisely the definition of span provided in (5). For this reason, the empirical coverage of the spanning hypothesis is similar to the empirical coverage of the X^0 hypothesis.

The spanning hypothesis is weaker in one respect, in that a span may be spelled out in a low position, effectively allowing heads to lower. Lowering is normally assumed to be restricted in the X^0 hypothesis. If there were a principled restriction on lowering that could be captured on the X^0 hypothesis and not on the spanning hypothesis, that would be an argument in favor of the X^0 hypothesis. However, there are several indications that the most commonly assumed restrictions on head-lowering are empirically inadequate.

For example, English finiteness in T lowers to V (Bobaljik 1995), which is supposed to occur under adjacency. But on a fine-grained analysis of the extended projection of the verb, the finite features are not located in a head which is adjacent to V; several heads such as nonfinite tense, aspect, and voice intervene. In particular, there are several arguments for an Asp projection in English between T and Voice, see Ramchand & Svenonius (2014)

for discussion and references.

It is also striking that adverbs do not prevent T lowering to V. Cinque (1999) has argued that the relatively strict ordering of adverbs suggests that they are associated with distinct projections, which further motivates heads between finiteness and the verb.

There are additional arguments for inflected heads being spelled out low in other languages, for example Haider (2000) for German and Adger et al. (2009) for Kiowa.

In the following subsections I discuss additional arguments in favor of the spanning hypothesis over the X^0 hypothesis.

3.1 Phonologically and lexically driven periphrasis

A well known example of periphrasis is seen in the English comparative, and it has been extensively discussed in the DM framework (Embick & Noyer 2001, Embick 2007, Embick & Marantz 2008, Bobaljik 2012). A pertinent aspect of the English comparative is that the alternation between the synthetic (*smarter*) and the analytic or periphrastic alternative (*more intelligent*) is affected by root information only available after lexical matching, since the synthetic comparative formation with *-er* is only possible with 'short' adjectives.

On the model adopted here, as well as in DM, lexical matching occurs after syntactic operations, so the difference between *smarter* and *more intelligent*cannot be due to syntax. On the spanning hypothesis, then, it must be possible for the span consisting of the comparative head and the adjective to be spelled out by one word or by two, without the syntactic structure bearing an additional w or @ feature in the case of 'long' adjectives.

The simplest way to state this is that the span consisting of, say, A-*a*-COMP, contains exactly one w and one @ feature and hence is spelled out as one word, all else being equal, and as expected exponent of the higher COMP head surfaces as a suffix. Suppletive forms like *worse* are simply portmanteaux for the whole A-*a*-COMP span.

However, the comparative exponent *-er*bears a contextual restriction which limits it to combining with short adjectives. In the case of a phonologically inappropriate host, such as *intelligent*, *-er*cannot spell out COMP. Instead, MORE surfaces, and is not suffixal. In fact, it is not even a reduced function word: it cannot surface with a reduced vowel, like the functional preposition

for, which can surface as [fər], but requires a full vowel as if it were a lexical word.

There are at least two plausible explanations for why *more*has the prosody of a lexical word, on the spanning hypothesis. One is that it is lexically specified with prosodic structure. Lexical specification of prosodic structure for certain elements is standardly assumed in phonology. Another is that *mo-* is a root (presumably a bound allomorph of *much*) which is inserted to 'rescue' the stranded affix *-er* (cf. Corver 1997), and the phonology of English assigns prosodic word structure to grammatical words which include roots.

On the X^0 hypothesis of word structure, however, there is a bigger problem. Forms like *smarter*are ordinary words, so they should be complex X^0s, formed by head movement. But if they were, then this X^0 complex would have to be undone in the case of *more intelligent*. To allow postsyntactic morphology to disassemble words in this fashion would to be to award it too much power to disrupt the sentence structure and would overgenerate.

If, on the other hand, forms like *more intelligent* show that COMP and A are not included in a single complex X^0, then some other operation must bring them together in the case of *smarter*. Embick (2007) posits a rule of Local Dislocation which has the form in (35).

(35) $\text{Deg[CMPR,SUP]}^\frown[\ldots \text{X} \ldots]_a \rightarrow [\ldots \text{X} \ldots]_a \oplus \text{Deg[CMPR,SUP]}$
Where the phonological form of $[\ldots \text{X} \ldots]_a$ meets the prosodic conditions

The tie bar indicates string adjacency, [CMPR,SUP] indicates that this rule applies both to the comparative and the superlative, and the circled plus sign indicates word-formation; so the rule states that if a Deg head with comparative or superlative features is adjacent to an adjective, then the the Deg head is suffixed to the adjective provided it meets the prosodic conditions.

There are several problems with this proposal. One is that it allows an unwanted mixing of syntactic and phonological information. In the two-stage lexical insertion model of Bye & Svenonius (2012), once the phonological exponents have been matched, no further reference can be made to syntactic features such as *a*.

If the condition on Local Dislocation is simply adjacency, then it is unclear why the comparative suffix cannot surface on material other than the head of the adjective phrase. For example, *filthy* and *pretty* allow comparative

formation, *filthier* and *prettier*, but not when modifying adjectives.

(36) a. The DuPonts are more filthy rich than the Smiths.
b. *The DuPonts are filthier rich than the Smiths.

Similarly, *light* allows the comparative form *lighter*, but only when it is lightness that it being compared, as in (37b); when a modified adjective like *light blue* is made comparative, *-er* cannot attach to the adjacent modifier *light*, so that (37b) cannot mean what (37a) means.

(37) a. This grass is much more light blue than what we've previously grown.
b. This grass is much lighter blue than what we've previously grown.

Thus, string adjacency is not the right condition for synthetic comparative formation. Modified adjectives like *filthy rich* arguably contain more structure than A-*a*, for example a Degree head with a null exponent. If the comparative head COMP is distinct from the Degree head Deg, and dominates it, then it is unsurprising that the highly selective affix *-er* which is an exponent of COMP cannot attach to Deg, and the periphrastic comparative is required in the case of *filthy rich*.

Another problem with the rule in (35) concerns suppletive and irregular forms. For example, the comparative of *bad* is the suppletive *worse*. The rule in (35) has to apply after lexical insertion, so that the prosodic conditions can be taken into consideration. But that means that the root *bad* should be present when the rule applies, apparently requiring it to be overwritten by *worse*.

In general, the rules of Local Dislocation posited in Embick & Noyer (2001), Embick (2007), and other works are not motivated by conceptual concerns, but are responses to empirical problems for the X^0 theory of word formation. Only unprincipled stipulations prevent them from overgenerating, for example allowing reordering of a verb and the determiner of its direct object, for a lexically specified class of cases. Such cases are unattested, to my knowledge.

3.2 Head movement

The X^0 theory of word formation partly unifies word formation with the theory of movement, which would be a positive reduction if it could be made to work, since many aspects of the theory of movement are well supported and relatively well understood, especially unbounded dependency formation (A-bar movement).

For example, there is a broad consensus that A-bar movement is cyclic, obeys some version of the extension condition, affects scope relations, and can be undone or disregarded in the interpretation (reconstruction effects). Head movement, on the other hand, seems irrepressibly different; it is apparently counter-cyclic, sometimes lowers features rather than raising them, violates the extension condition, has no effect on scope, and shows no reconstruction alternations (if anything, it obligatorily reconstructs, as if it had never happened).[12]

As an illustration of the apparent counter-cyclicity of head movement, it is unclear how Mainland Scandinavian languages can have V-to-C without V-to-T. The Norwegian verb, for example, does not independently move to T (as seen in embedded clauses, where it follows medial adverbs in standard Norwegian).

(38) a. Visstnok fant Octavia aldri den røde sokken.
apparently found Octavia never the red sock.DEF
'Apparently, Octavia never found the red sock'
b. Jeg tror at Octavia aldri fant den røde sokken.
I believe that Octavia never found the red sock.DEF
'I believe that Octavia never found the red sock'

Since head movement is strictly local (Travis 1984), attraction by C in main clauses should attract only T, as in English, perhaps triggering do-support. In fact, auxiliaries do move independently of the lexical verb, as in English.

(39) Visstnok har Octavia aldri funnet den røde sokken.
apparently has Octavia never found the red sock
'Apparently Octavia has never found the red sock'

But in Norwegian, in the absence of an auxiliary, the main verb moves to C,

[12]See Platzack (2013) and references there for discussion of the dearth of syntactic properties of head movement.

as seen in (38a). So at the point where T is merged, the verb must not move, because the clause might be an embedded clause. But once main clause C is merged, the verb must countercyclically move through T.

In the spanning hypothesis, following Brody's Mirror Theory, there can be no head movement as there is no distinction between heads and phrases. What replaces head movement is variable linearization of heads, controlled by the @ feature. Since linearization is a matter of phonological expression, occurring late in the derivation, no countercyclicity is introduced into the syntax (relinearization of V-v to C might be dependent on V-v being at a phase edge, as in Fox & Pesetsky 2005).

4 Conclusion

It is a tenet of DM that terminal nodes are the locus of lexical insertion (see e.g. Embick & Noyer 2001, Embick & Marantz 2008).

I have suggested instead in (7) that Lexical Matching, and hence insertion, must apply to spans.

The difference is partly a matter of granularity. Heads are traditionally viewed as bundles of features. However, there has been a steady development over the past thirty-five years to recognize increasing numbers of features as heads in their own right. Thus, a single head on a conservative analysis (INFL, for example, in Chomsky 1981) might correspond to multiple heads on a more progressive analysis (Agr and T, in Pollock 1989).

The extreme point of this development is to posit a single syntactically active feature per head (Kayne 2005). Bobaljik (2012), for example, argues that the superlative form of an adjective is built on the comparative, and that cross-linguistically valid restrictions on syncretism cannot be explained if languages are free to posit a single head which combines the two meaning components. He observes that this would follow from an atomicity hypothesis like that suggested by Kayne.

If anything like atomicity is true, then there are far more heads in the syntax than was countenanced in the 1980's (cf. Cinque 1999), and far more portmanteaux in the morphology (cf. Starke 2009). Portmanteaux are straightforwardly handled on the spanning hypothesis, whereas on the X^0 hypothesis, they require additional mechanisms such as Fusion.

Essentially, these developments move a burden from the theory of feature geometries (see e.g. Harley & Ritter 2002) to the theory of extended

projections (see Adger 2013, Ramchand & Svenonius 2014).

The word has central relevance for syntax, morphology, and phonology, and perhaps even semantics, yet there is as of yet no well-developed theory of its source. The phase-based theory of the word (Marantz 2001; 2007, Marvin 2002, Newell 2008) is in this regard something of a breakthrough, but still relies on central features of DM, terminal node insertion, and the X^0 theory of word formation.

In this paper I have argued that the X^0 hypothesis of words leads to certain impasses, and that thinking of words in terms of spans may help us move beyond them.

I have also argued that the spanning hypothesis of word structure allows for a more minimalist theory of morphology. In many individual analyses, it obviates mechanisms such as Local Dislocation, Fusion, Morphological Merger, and Head Movement which are invoked in competing analyses, opening up the possibility that such mechanisms could be dispensed with altogether.

References

Adger, David, 2007. Stress and phasal syntax. *Linguistic Analysis* 33(3-4):238–266.

Adger, David, 2013. *A Syntax of Substance*. Cambridge, Ma.: MIT Press.

Adger, David, Harbour, Daniel & Watkins, Laurel J., 2009. *Mirrors and Microparameters: Phrase Structure Beyond Free Word Order*. Cambridge: Cambridge University Press.

Baker, Mark C., 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Baker, Mark C., 1996. *The Polysynthesis Parameter*. Oxford Studies in Comparative Syntax. New York: Oxford University Press.

Bermúdez-Otero, Ricardo, 2011. Cyclicity. In *The Blackwell Companion to Phonology*, volume 4, Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume & Keren Rice, eds., 2019–2048. Oxford: Blackwell.

Bobaljik, Jonathan David, 1995. *Morphosyntax: The Syntax of Verbal Inflection*. Ph.D. thesis, MIT, Cambridge, Ma.

Bobaljik, Jonathan David, 2012. *Universals in Comparative Morphology: Suppletion, Superlatives, and the Structure of Words*. Cambridge, Ma.: MIT Press.

Borer, Hagit, 2005. *In Name Only (Structuring Sense, vol. I)*. Oxford: Oxford University Press.

Brody, Michael, 2000a. Mirror Theory: Syntactic representation in Perfect Syntax. *Linguistic Inquiry* 31(1):29–56.

Brody, Michael, 2000b. Word order, restructuring, and Mirror Theory. In *The Derivation of VO and OV*, Peter Svenonius, ed., 27–43. Amsterdam: John Benjamins.

Bye, Patrik & Svenonius, Peter, 2012. Non-concatenative morphology as epiphenomenon. In *The Morphology and Phonology of Exponence*, Jochen Trommer, ed., 427–495. Oxford: Oxford University Press.

Caha, Pavel, 2009. *The Nanosyntax of Case*. Ph.D. thesis, University of Tromsø.

Cardinaletti, Anna & Starke, Michal, 1999. The typology of structural deficiency: A case study of three classes of pronouns. In *Clitics in the Languages of Europe*, Henk van Riemsdijk, ed., 145–233. Berlin: Mouton.

Carlson, Greg, 2006. The meaningful bounds of incorporation. In *Non-definiteness and Plurality*, Svetlana Vogeleer & Liliane Tasmowski, eds., 35–50. Amsterdam: John Benjamins.

Chomsky, Noam, 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam, 2001. Derivation by phase. In *Ken Hale: A Life in Language*, Michael Kenstowicz, ed., 1–52. Cambridge, Ma.: MIT Press.

Chomsky, Noam, 2008. On phases. In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta, eds., 133–166. Cambridge, Ma.: MIT Press.

Cinque, Guglielmo, 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York: Oxford University Press.

Corver, Norbert, 1997. Much-support as a last resort. *Linguistic Inquiry* 28:119–164.

Déchaine, Rose-Marie, 2015. What "spell-out" reveals: Niger-Congo prosodification constrains the syntax-semantics interface. In *Current Research in African Linguistics: Papers in Honor of Ọladele Awobuluyi*, Ọlanikẹ Ọla Orie, Johnson Fọlọrunṣo Ilọri & Lendzemo Constantine Yuka, eds., 287–352. Newcastle: Cambridge Scholars.

Di Sciullo, Anna-Maria & Williams, Edwin, 1987. *On the Definition of Word*. Linguistic Inquiry Monographs; 14. Cambridge, Ma.: MIT Press.

Dryer, Matthew, 1989. Plural words. *Linguistics* 27:865–895.

Embick, David, 2007. Blocking effects and analytic/synthetic alternations. *Natural Language & Linguistic Theory* 25(1):1–37.

Embick, David, 2010. *Localism versus Globalism in Morphology and Phonology*. Cambridge, Ma.: MIT Press.

Embick, David & Marantz, Alec, 2008. Architecture and blocking. *Linguistic Inquiry* 39(1):1–53.

Embick, David & Noyer, Rolf, 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4):555–595.

Fox, Danny & Pesetsky, David, 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31(1-2):1–45.

Haider, Hubert, 2000. OV is more basic than VO. In *The Derivation of VO and OV*, Peter Svenonius, ed., 45–67. Amsterdam: John Benjamins.

Halle, Morris & Marantz, Alec, 1993. Distributed Morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Kenneth Hale & Samuel Jay Keyser, eds., 111–176. Cambridge, Ma.: MIT Press.

Harley, Heidi & Ritter, Elizabeth, 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3):482–526.

Ito, Junko & Mester, Armin, 2012. Recursive prosodic phrasing in Japanese. In *Prosody matters: Essays in honor of Elisabeth Selkirk*, Toni Borowsky, Shigeto Sugahara, Mariko Sugahara & Takahito Shinya, eds., 280–303. Sheffield: Equinox.

Kahnemuyipour, Arsalan, 2009. *The Syntax of Sentential Stress*. Oxford: Oxford University Press.

Kayne, Richard S., 1994. *The Antisymmetry of Syntax*. Cambridge, Ma.: MIT Press.

Kayne, Richard S., 2005. Some notes on comparative syntax, with special reference to English and French. In *The Oxford Handbook of Comparative Syntax*, Guglielmo Cinque & Richard S. Kayne, eds., 3–69. New York: Oxford.

Kornfilt, Jaklin, 1997. *Turkish*. London: Routledge.

Kratzer, Angelika & Selkirk, Elisabeth, 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24(2-3):93–135.

Marantz, Alec, 2001. Words. Ms. MIT.

Marantz, Alec, 2007. Phases and words. In *Phases in the Theory of Grammar*, Sook-hee Choe, ed., 196–226. Seoul: Dong-in.

Marantz, Alec, 2013. Locality domains for contextual allomorphy across the interfaces. In *Distributed Morphology Today: Morphemes for Morris Halle*, Ora Matushansky & Alec Marantz, eds., 95–115. Cambridge, Ma.: MIT Press.

Marvin, Tatjana, 2002. *Topics in the Stress and Syntax of Words*. Ph.D. thesis, MIT, Cambridge, Ma.

Nespor, Marina, Shukla, Mohinish, van de Vijver, Ruben, Avesani, Cinzia, Schraudolf, Hanna & Donati, Caterina, 2008. Different phrasal prominence realizations in VO and OV languages. *Lingue e linguaggio* 7(2):139–168.

Newell, Heather, 2008. *Aspects of the Morphology and Phonology of Phases*. Ph.D. thesis, McGill University, Montréal.

Nikolaeva, Irina, 2013. Periphrasis in Tundra Nenets. In *Periphrasis: The Role of Syntax and Morphology in Paradigms*, Marina Chumakina & Greville G. Corbett, eds., 77–140. Oxford: The British Academy.

Noyer, Rolf, 1998. Vietnamese 'morphology' and the definition of word. *University of Pennsylvania Working Papers in Linguistics* 5(2):65–89.

Platzack, Christer, 2013. Head movement as a phonological operation. In *Diagnosing Syntax*, Lisa Lai-Shen Cheng & Norbert Corver, eds., 21–43. Oxford: Oxford University Press.

Pollock, Jean Yves, 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20(3):365–424.

Postal, Paul, 1969. Anaphoric islands. In *Chicago Linguistic Society*, volume 5, 205–239.

Ramchand, Gillian, 2008. *Verb Meaning and the Lexicon*. Cambridge: Cambridge University Press.

Ramchand, Gillian & Svenonius, Peter, 2014. Deriving the functional hierarchy. *Language Sciences* 46(B):152–174.

Rijkhoff, J., 2002. *The Noun Phrase*. Oxford: Oxford University Press.

Roberts, Ian, 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge, Ma.: MIT Press.

Scheer, Tobias, 2008. Why the prosodic hierarchy is a diacritic and why the interface must be direct. In *Sounds of Silence: Empty Elements in Syntax and Phonology*, Jutta Hartmann, Veronika Hegedus & Henk van Riemsdijk, eds., 145–192. Oxford: Elsevier.

Selkirk, Elisabeth, 1996. The prosodic structure of function words. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, James L. Morgan & Katherine Demuth, eds., 187–213. Hillsdale, NJ: Lawrence Erlbaum.

Selkirk, Elisabeth, 2011. The syntax-phonology interface. In *The Handbook of Phonological Theory*, John Goldsmith, Jason Riggle & Alan Yu, eds. Oxford: Blackwell, 2nd edition.

Siddiqi, Daniel, 2009. *Syntax within the Word: Economy, Allomorphy, and Argument Selection in Distributed Morphology.* John Benjamins.

Son, Minjeong & Svenonius, Peter, 2008. Microparameters of cross-linguistic variation: Directed motion and resultatives. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, Natasha Abner & Jason Bishop, eds., 388–396. Somerville, Ma.: Cascadilla.

Starke, Michal, 2009. Nanosyntax: A short primer to a new approach to language. In *Tromsø Working Papers on Language and Linguistics: Nordlyd 36.1,* Special issue on Nanosyntax, Peter Svenonius, Gillian Ramchand, Michal Starke & Knut Tarald Taraldsen, eds., 1–6. Tromsø: University of Tromsø. Available at http://septentrio.uit.no/index.php/nordlyd.

Šurkalović, Dragana, 2013. Modularity, phase-phase faithfulness and prosodification of function words in English. *Nordlyd* 40(1):301–322.

Svenonius, Peter, 2005. Extending the extension condition to discontinuous idioms. *Linguistic Variation Yearbook* 5:227–263.

Svenonius, Peter, 2012. Spanning. Ms. University of Tromsø, available at ling.auf.net/lingBuzz/001501.

Svenonius, Peter, 2016. Complex predicates as complementation structures. In *Approaches to Complex Predicates*, Léa Nash & Pollet Samvelian, eds., 212–247. Leiden: Koninklijke Brill.

Svenonius, Peter, Ramchand, Gillian, Starke, Michal & Taraldsen, Knut Tarald (eds.), 2009. *Tromsø Working Papers on Language and Linguistics: Nordlyd 36.1,* Special issue on Nanosyntax. Tromsø: University of Tromsø. Available at http://septentrio.uit.no/index.php/nordlyd.

Travis, Lisa, 1984. *Parameters and Effects of Word Order Variation.* Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Ma.

Williams, Edwin, 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12(2):245–274.

Woodbury, Anthony C., 1987. Meaningful phonological processes: A consideration of Central Alaskan Yupik Eskimo prosody. *Language* 63(4):685–740.