A span is a thing: A span-based theory of words*

Peter Svenonius

CASTL, University of Tromsø – The Arctic University of Norway

1. Definitions

This paper contains a modest claim swathed in a more ambitious proposition. Modestly, I argue that **a span is a thing**: namely, that a parsimonious description of grammar in the current state of syntactic theory will include reference to spans. More ambitiously, I promote a theory of words based on spans, developed in part elsewhere (especially Bye and Svenonius 2012, Svenonius 2012, and Svenonius 2016b). Clearly, then, it is incumbent on me to define both spans and words.

1.1 Spans

To define a span is easy, if you grant the premise (see Svenonius 2016a) that there is a complement relation (R_C) that is distinct from the relations relevant for specifiers ('subjects' or arguments of predicates) and for adjuncts (modifiers of predicates). Importantly for the concept of span, the heads in an extended projection (Grimshaw 2005) are related by the complement relation. Derivational morphology is also related to its base by complementation. Internal arguments, on the other hand, are often specifiers.

In the definition in (1), the span is represented with the bottom to the left, top to the right (e.g., $\langle$V, Voice$\rangle$ for V=h_1, Voice=h_2, where V is the head of the complement of Voice).

(1) a. A span is a head (a minimal X^0) or a sequence of heads $\langle h_1, h_2 \ldots \rangle$ such that for each $n>1$, $R_C(h_{n-1},h_n)$ (Svenonius 2012, Merchant 2015)
b. $R_C(h_m,h_n) = h_m$ is the head of the complement of h_n

The term EXTENDED PROJECTION SPAN or EP-span refers to spans which include only material from a single extended projection. A MAXIMAL SPAN is a span which is not the complement of any head.

*Thanks to the audience at NELS 50 at MIT for discussion, and to David Adger for comments on a draft, and to the editors for their hard work.

1.2 Words

1.2.1 Syntactic words

To define a word will require more work. There are at least two distinct senses of word that are relevant for syntax. In the one sense, a compound like *cheesecake* is a word, and in the other sense, *cheesecake* is made up of two words. I will use SYNTACTIC WORD for syntactically free elements like *cheesecake*.

A phonological word like *say's* ([seɪz]) in *Who did you say's coming to the party?* consists of two syntactic words; *say* is free in the sense that it can be preceded or followed by material other than that spelling out its own extended projection, and so is *'s*, a phonologically reduced form of *is*. In §4.2, I will argue that the irregular stem change in *says* [sɛz] is only possible because the stem and the suffix *-s* are in the same word and the same span.

The Igloolik (Inuit) sentence meaning 'he/she made a big house' is a single word, according to Compton and Pittman 2010.

(2) iglu-jjua-liu-lauq-tuq
house-big-make-PAST-DEC3SG
'He/She made a big house'

Compton and Pittman argue that every (finite) CP and (case-marked) DP phase is a word in Igloolik, where a phase is a unit of cyclic transfer to the semantic and phonological interfaces (Chomsky 2004). In (2), the subject is null and the object is incorporated. In English, a sentence like *She built a big house* would be spelled out in as many as three separate phases: *She* and *a big house* are case-marked noun phrases and so plausibly phases (Svenonius 2004), and *built* would be spelled out in a separate round as the content of one of the phases in the clause, possibly the vP.

Furthermore, a noun phrase like *a big house*, though plausibly a single phase, spells out as several words. I suggest that isolating languages like English differ from languages like Igloolik in assigning isolable interface interpretations to certain units below the phase level. Those units are syntactic words.

Just as a phase has a phase head, a sub-phase unit for interface interpretation must have a property distinguishing it from arrangements of heads which are not syntactic words. I will call that distinction a syntactic word label (Svenonius 2016b). Following Brody 2000, I use the symbol @ for the syntactic word label (pronounced like the preposition *at*).

In the English sentence *She built a big house*, there are five word labels, so there are five nodes bearing @. In the corresponding Igloolik sentence in (2), there is one.

If every maximal span (including every specifier and adjunct) constitutes a cycle of spell-out (as in Uriagereka 1999), then all previously unspelled-out material in each cycle constitutes a span. However, examples like (2) show that some specifiers and adjuncts do not spell out independently. I call specifiers and adjuncts which do not spell out as syntactic words DEFECTIVE: they are spans which do not themselves bear syntactic word labels. Defective specifiers include clitic pronouns, and defective adjuncts include incorporated modifiers (like *-jjua-* 'big' in (2)).

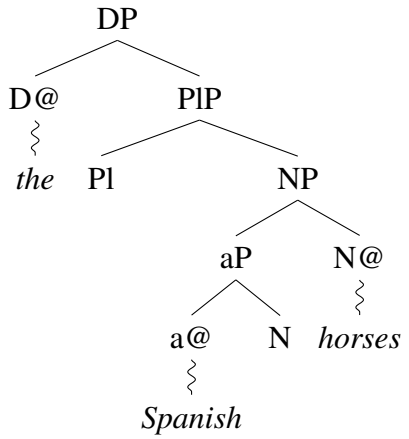
Now the syntactic word can be defined as in (3) (see also Svenonius 2016b).

(3) Syntactic word: The smallest unit assigned an isolable interpretation at both interfaces. An @-span, which is a span containing exactly one @, along with any defective specifiers and adjuncts which are dependent on that span.

The definition in (3) does not resolve how to delimit the @-spans in a span which contains more than one @.

Consider an English noun phrase like *the Spanish horses*, as illustrated in (4), where the definite article is D, the word *Spanish* a span ⟨N, a⟩, and the plural-marked noun a span ⟨N, Pl⟩. Nationality adjectives like *Spanish* have been argued to be structurally lower than the plural (Svenonius 2008b and references there).

(4)



There are three syntactic words here, so there are three nodes marked with @. The definite article linearizes in D, so there is an @ there, and the noun linearizes in N, together with the Pl suffix, so there is an @ in N (as if Pl lowers to N). In the adjectival span, there is only one word so we can't immediately tell in which of the two heads the syntactic word linearizes. I have assumed that @ is on the derivational head a (as if there were head movement from N to a).

Here we see here that the material in between the two @'s belongs with the lower head. That principle works generally for English, but I do not exclude the possibility of parametrization.

1.2.2 Stem-words

A sense of word distinct from the syntactic word is deployed when we say that a compound is made up of two words, or that there is a word boundary internal to a complex syntactic word. I will coin the term STEM-WORD to allow me to characterize these kinds of situations. I argue in Svenonius 2016b that the distinctive properties of stem-words are the result of a cycle of lexical insertion, similar to what has been proposed in Marvin 2002, Marantz 2007, and Newell 2008, but crucially for me, this cycle is prior to Chomsky's (2004) phase-level Transfer to the interfaces.

This stage of lexical access can be compared to Level 1 of Kiparsky's (1982) Level-ordered morphology, or to the stem cycle of Bermúdez-Otero 2017. It is induced by cyclic nodes, as proposed in Marantz 2007. It allows the production of a set of candidate words which can be compared to stored lexical entries (as discussed by Bermúdez-Otero 2012).

In Svenonius 2016b I introduce a property w borne by certain heads which induces a round of lexical insertion.

The clearest cases of word-internal word boundaries can be attributed to an element that contains a root in the sense of Marantz 1997 or Borer 2005.

(5) Stem-word: The output of a cycle of lexical insertion which includes a root; a w-span, a span in which the top node bears w

The stem-word corresponds approximately to the prosodic word in phonology (modulo postsyntactic adjustments in the phonology).

What is commonly called a lexical word (like *word*) is then a syntactic word in the sense of (3) which contains a stem-word in the sense of (5), and what is commonly called a function word (like *a*) is a syntactic word which does not contain a stem-word.

In a compound like *blackbird*, one stem-word is adjoined to another. In work in progress (Svenonius 2018), I examine differences between compounds, which are not separated by 'head movement' (which I analyze without movement), from clusters, which can be separated, for instance under V2 in German. I suggest that the difference can be related to the difference between adjunction and complementation.

In a verb cluster, multiple w-spans contained in a single @-span linearize together. V2 in Germanic languages involves a higher @ in C, hence a second @-span, which provides a distinct point of linearization for the highest w-span in the cluster.

Phonological and syntactic diagnostics sometimes fail to distinguish a cluster from a compound, in which case semantics may be the only guide. In any event, on the spanning account, no movement is needed in either case to combine stem-words to each other.

2. We need a syntactic theory of word formation

A syntactic theory of words is clearly needed, since phonology alone cannot predict the boundaries of phonological words.

For example, a German separable prefix verb like *aufführen* 'perform' can be a single word or two, depending on whether it is the finite verb in a matrix clause. If it is, @ in C causes it to separate syntactically into two words (*Sie führen das Stück auf* 'They put the play on'). We can tell there are two syntactic words in that case because they can be separated by specifiers and adjuncts to the extended projection of the verb.

When it is not separated, for example in an embedded clause, *aufführen* forms a cluster and takes on certain properties of single words, for example it can combine with a prefix like *ur-* 'original' (*ob sie das Stück uraufführen*, 'if they perform the play for the first time').

The prefixed cluster *uraufführen* 'perform for the first time' cannot be the finite verb of a main clause in German (Vikner 2005 and references there). I argue in Svenonius 2017b that this is not due to any morphological or phonological ill-formedness. Rather, it is because syntactic words are assigned interpretations at Transfer, and only the unseparated word *aufführen* has the right semantic content to serve as the base of *ur-*.

In a main clause, the verb *uraufführen* cannot be left in situ, because @ in C in German main clauses causes a syntactic word to linearize there. The whole cluster cannot be spelled out in the second position, because an internal stem-word juncture in separable prefix clusters causes them to strand their separable prefix in the VP when the finite verb word is

attracted to second position. The V2 linearization results in two smaller words, neither of which can be interpreted semantically with *ur-* prefixed. At LF, the two parts of a simple separable verb like *aufführen* can be united and their idiomatic interpretation is retrieved, but not if *ur-* has already been introduced into the meaning of either part.

The conjecture expressed in (3) that words have interpretations at both interfaces plays a crucial role here. It is the isolable interpretability of the individual parts, *führen* and *auf*, which allows them to be stored and reconstituted, much as with the 'idiomatically combining expressions' of Nunberg et al. 1994.

The syntactic word conjecture also has other implications. For example, taken at face value, it should entail that no word completely lacks an isolable interpretation. There are syntactic heads such as Topic heads which have an influence on semantic and phonological interpretation, but which do not correspond to an isolable unit of form or interpretation. The prediction here is that such a head will not by itself be a syntactic word; it can at most be a bound part of a syntactic word (or be manifested in the order or prosody of other words). Similarly, an agreement probe will not by itself be a syntactic word. Of course, words may have different senses in different contexts. The words used as expletives have interpretations, even if the interpretations are not used when those words are deployed as expletives.

Under German V2, the finite verb *führen* and the stranded particle *auf* are arguably spelled out in different phases. This means that the meaning of *aufführen* can be reconstituted from the interpretation of *auf* in one phase and the interpretation of *führen* in another. One might be forgiven for thinking that the argument from *ur-* shows only that each phase receives an interpretation. I am making a stronger claim, that individual words within a phase receive interpretations.

This can be illustrated with a parallel argument from English. The two words comprising particle verbs like *turn over* are spelled out in the same phase (for example the VP can be fronted in *Turn them over, I did*). Yet the English prefix *re-* shows the same kind of restriction as German *ur-* (Williams 2014): a word like *overturn* provides the right semantic base for *re-*, allowing *reoverturn*. A particle verb like *turn over* also has the right semantics, but only distributed over two words. As a result, *re-* cannot attach (**return over*). On my account, this is a direct result of syntactic words being assigned interpretations at spell-out. *Return* gets the wrong meaning for composition with *over*, and access to *turn over* comes too late. Similarly, *reinput*, *reupload*, but **reput in*, **reload up*, and so on.

3. Movement is not working for word formation

In standard DM, the heavy lifting in word formation is performed by head movement, constrained by the Head Movement Constraint (HMC, Travis 1984), which limits head movement to spans (call this the spanmate condition). The biggest differences between head movement and spelling out spans directly as words are (i) that head movement is supposed to be unified with movement, and to the extent that it is, is conceptually motivated in a way that span spell-out is not, and (ii), because of (i), head movement predicts that complex words will be linearized in the position of their highest affix, while span spell-out makes no such prediction.

However, the adoption into relatived minimality of a head-phrase distinction (Rizzi 1990) as a basic property to which locality is sensitive is fundamentally a stipulation which has not been independently justified. Without it, the HMC is not derived and so head movement and phrasal movement appear to have distinct kinds of constraints on their locality. Phrasal movement can be stated in terms of classes of attracting features, but this does not derive a spanmate condition.

Furthermore, the prediction that complex words linearize high fails. There are many cases in which an affix appears to lower, as with English T-to-V, which is problematic for a movement analysis but is perfectly consonant with a spanning approach (a V-T span may linearize in V, Brody 2000). See Adger et al. 2009 for a detailed case from Kiowa.

Consider Icelandic, in which the noun stem bears up to three inflectional suffixes expressing gender, number, definiteness, and case, which can be analyzed as N-Num-D-K, as in *hest-ar-n-ir* "horse-M.SG.NOM-DEF-M.SG.NOM", 'the horses (NOM)'. The prediction of head movement would be that the N moves to Num, then D, then K. But adjectives still precede the noun, as in *íslensku hestarnir* 'the Icelandic horses (NOM)', so head movement to D and to K has apparently not occurred.

Numerals follow the definite-marked noun (*hestarnir fjórir* 'the four horses'), which has led some to postulate phrasal movement of [A N] across the numeral and D (e.g., Julien 2005). But as Harðarson (2017) shows, this fails to capture the constituency of Icelandic noun phrases, where NP-internal material remains postnominal and there is no evidence for an [A N] syntactic constituent. Harðarson suggests that N undergoes head movement across the numeral, and that adjectives which precede the definite noun are focus-fronted. Another possibility is that the fronting of adjectives is related to cluster formation, for example if weak inflected adjectives lose @ when dominated by @ in the nominal EP span. A third alternative is that it is the numeral that clusters to the right of a noun low in its EP span.

4. Predictions: What spanning can do for you

The architecture outlined in §1, whereby a cycle of stem-word lexicalization precedes the spell-out of the phase, and in which syntactic words within a phase are interpreted in parallel, has a variety of applications and makes some predictions which can be tested.

4.1 Portmanteaux

Recall that both stem-words and syntactic words were defined in terms of spans. In general, exponents map to spans. I propose that this is because the syntactic information associated with an exponent is restricted to a list of NODE IDENTIFIERS, where a node identifier is a category (positionally defined), a set of syntactico-semantic feature structures, or a combination of the two. Since the list of node identifiers is one-dimensional, the exponent maps to a one-dimensional object like a span, not to an object with internal branching.

Some of the earliest applications for spans were in the analysis of portmanteaux (e.g., Abels and Muriungi 2008, Son and Svenonius 2008, Svenonius 2012). For example, *mice* could be the span of N and Pl, rather than a contextually conditioned allomorph of *mouse*

which reciprocally triggers a zero contextual allomorph of the plural, as on the DM analysis.

A spanning theory with portmanteaux can eliminate many of the zeros which must be posited in DM. However, there are contexts where a learner may be motivated to posit a zero affix. For example, for all lexical verbs in English, the base form is identical to the present tense when the subject is not third person singular. The generalization that every verb stem can spell out present tense T is parsimoniously captured as a zero exponent for T.

4.2 Conditioned allomorphy

An exponent can be said to 'lexicalize' a span of categories, for example when the exponent *-(e)d* lexicalizes the category T in the past. Conditioned allomorphy (CA) is a phenomenon whereby an allomorph A of a target is triggered or conditioned by something which is not a feature of the span that A lexicalizes, for example if the *-(e)n* allomorph of the perfect participle is chosen when the stem is a strong verb: the allomorph *-(e)n* lexicalizes Asp[past], but it is conditioned by a strong verb.

CA must be distinguished from INDEX AGREEMENT; ϕ features (person, number, and gender) and related features (case, honor), as well as tense, can be copied syntactically by probes, and the appearance of an exponent which is conditioned by such features does not require CA.

The standard way of treating CA, as a statement of the context of insertion for an allomorph, is powerful if the context is not constrained. The restriction of lexical insertion to the context of a span predicts a more limited distribution for CA.

Since a conjugation class is a lexical listing of verbs which submits to no semantic or phonological characterization, the distribution of the *-(e)n* suffix is a case of LEXICALLY CONDITIONED SUPPLETIVE ALLOMORPHY (LCSA), which must be distinguished from phonologically conditioned suppletive allomorphy (PCSA) and from syntactically conditioned suppletive allomorphy (SCSA).

4.2.1 PCSA

In PCSA, a phonological trigger conditions the insertion of a particular allomorph, for example the Dyirbal ergative suffix *-ŋgu* discussed in Bye 2007 which is attached to any disyllabic stem ending in a vowel (elsewhere, the suffix is *-nu* after a vowel, *-Cu* after a nasal or a palatal glide, and *-ɽu* after a liquid, Dixon 1972).

In keeping with its phonological nature, PCSA can be assumed to be subject to string adjacency (possibly on a tier). But PCSA has been observed to cross word boundaries, for example, the Welsh definite article (Hannahs and Tallerman 2006) and the English indefinite article.

It is interesting that the best examples of PCSA across a syntactic word boundary involve a target which is a free function word, as opposed to a bound affix. The cyclic theory of words outlined in §1 gives a possible reason, if this pattern should hold up across a broader survey. Recall that stem-words undergo at least one round of lexical inser-

tion. In the maximal span, there may be additional function words outside the stem-word, 'stranded' as it were. The Welsh and English articles are stranded in this way; they are not subject to lexical insertion until the maximal span spells out.

If the cycle of lexical insertion at the stem-word level resolves competition among allomorphs, then allomorphs are fixed within the stem-word before material outside the stem-word becomes visible (setting aside domain expansion for the present discussion). In that case, PCSA will never allow a trigger outside a stem-word to condition a target inside, explaining why transword PCSA is only ever observed to target stranded function words.

4.2.2 SCSA

SCSA can be illustrated with the English verb system, when a form like *broke* is selected as the allomorph of *break* in the participle (setting aside an alternative treatment in terms of index agreement copying tense values to v, as suggested by Adger 2003). Here, the conditioning factor is a syntacticosemantic feature which I will call [past].

SCSA, unlike PCSA, is not subject to string adjacency. For example, Merchant 2015 argues that a perfective head can be part of the condition for a suppletive form of the Greek verb stem across an overt Voice head. Bonet and Harbour 2012 illustrates nonadjacent allomorphy in Kiowa inflection triggered by the presence of a transitivity head. Michael 2014 discusses nonadjacent allomorphy in Nanti inflection for the realis/irrealis distinction.

Merchant proposes that the locus of SCSA is the span. In all of the clear examples of which I am aware, it is also confined to the syntactic word. A possible explanation for these apparent restrictions on SCSA would be that SCSA is a subcase of the same central mechanism that is involved in spell-out, the association of node identifiers (see §4.1) in an exponent with nodes in a syntactic span. If the node identifiers include category information, the exponent lexicalizes the node. if the node identifier contains only featural information, the feature conditions the exponents by SCSA.

Categories are positionally defined (e.g., a category Asp2 may be defined as being located above Asp1 and below T in an extended projection), while features may appear in more than one category (e.g., a feature [past] may appear in T in the past but in Asp in the perfect). Each category is lexicalized by a single exponent (with portmanteau exponents lexicalizing more than one category), but, I suggest, an individual *feature* may be linked to more than one exponent.

Take for example the participle *broke-n*, where there are two exponents, lexicalizing at least two syntactic heads. The suffix *-n* lexicalizes at least the Asp head and is restricted to the perfect (setting aside the relation between the perfect and the passive). So the node identifier for *-n* is at least ⟨Asp[past]⟩. *Broke* lexicalizes at least V and is conditioned by the [past] feature, whether on Asp or on T. In that case [past] is one of the node identifiers associated with the exponent *broke*; something like ⟨V_{BREAK}, [past]⟩. Note that the second node identifier is a feature only, with no category.

Both *broke* and *-n* spell out [past], but only *broke* spells out V, and only *-n* spells out Asp, so neither can be omitted. **Broke-d* would also spell out the root and Asp[past], but the LCSA of *-(e)n* would block *-(e)d*. **Break-en* would also spell out the categories and

features, but in the competition among allomorphs, multiple exponence may be favored where available, leading *broke-n* to prevail.

This would entail that **broke-d* could also spell out *break* in the past. Possibly, the past suffix *-(e)d* is blocked by a null past tense suffix which is conditioned under LCSA by strong verbs (as on the DM analysis). For another compatible solution to the blocking pattern, see Adger 2014.

4.2.3 LCSA

The third type of conditioned allomorphy is LCSA, which I have already illustrated with the *-(e)n* participle ending. The verbs it combines with (strong verbs) cannot be defined semantically or phonologically, and so must be listed, which is what makes this LCSA.

LCSA, like PCSA, appears to be restricted to string adjacency (I am optimistic that the apparent counterexample from Nanti discussed by Inkelas (2017) can be reanalyzed as SCSA). But unlike PCSA, it is not observed to cross syntactic word boundaries. The usual treatment of LCSA does not distinguish it from PCSA or from SCSA, but on the account I have developed here, there is a way to derive its particular locality profile.

In Svenonius 2017a I proposed that declension class in Norwegian is pseudophonology: it is a feature which is (i) part of the phonological makeup of the exponent, but (ii) not a member of the phoneme inventory of the language, hence has no consistent phonetic interpretation.

The fact of (i) has the result that it can condition allomorphy in the first cycle of lexical insertion. But due to (ii), it is eliminated by the first cycle of phonological computation.

If generalized to LCSA more broadly, this has the result that LCSA is subject to adjacency, like PCSA, but cannot cross a syntactic word boundary, as desired.

5. An apparent challenge: Edge inflection

The account of conditioned allomorphy in §4.2 is more restrictive than what is usually assumed. If there is LCSA or SCSA which is sensitive to conditioning elements outside the span, then it is too restrictive and must be weakened. One apparent class of cases is those described as edge inflection, inflection at the edge of a phrase even when the edge is not a head in the extended projection.

Simple cases in which an exponent of a functional head appears at the edge of a phrase can be treated as clitics. Edge inflection is invoked when LCSA or SCSA of the exponent is sensitive to properties of the word to which it attaches even when that word is not part of the EP-span of the exponent. I illustrate with an example from North Sámi which has approximately the right structure.

In the North Sámi examples of illative noun phrases in (7) (from Nickel 1994), the demonstrative and numeral are marked with the illative case only if they not followed by a word inflected for illative. The numeral otherwise appears in the genitive, the demonstrative in a form which is syncretic for accusative and genitive, here glossed genitive. Nominative forms appear in (6) for comparison (NM for NOM[inative] in (6a) for reasons of space). See Svenonius 2008a for details of the morphology of case in North Sámi.

(6) a. ma-t golbma dálu
which-NM three.NM farm.NM
'which three farms'

b. ma-t golbma
which-NOM three.NOM
'which three'

c. ma-t
which-NOM
'which (one)'

(7) a. ma-n golmma dállu-i
which-GEN three.GEN farm-ILL
'to which three farms'

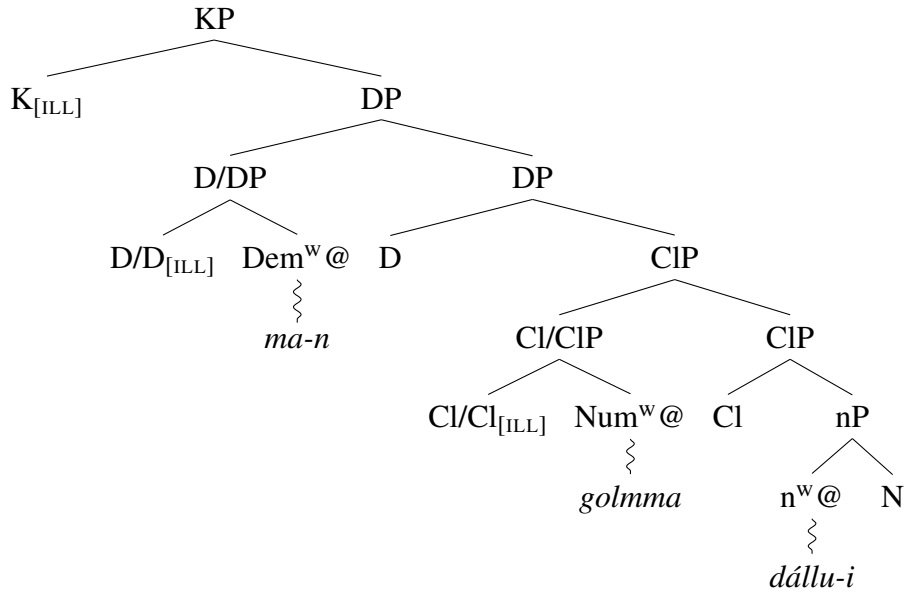
b. ma-n golbm-ii
which-GEN three-ILL
'to which three'

c. ma-sa
which-ILL
'to which (one)'

The result is that the North Sámi noun phrase appears to show illative case marking at its right edge, with allomorphy conditioned by the host word, since we see *-sa* on the demonstrative but *-(i)i* plus a trigger for a strong grade consonant center on the noun and numeral (leading to gemination of *-l-* in *dállui* in (7a) and glottalization of *-lmm-* in *golbmii* in (7b); there is also a strong grade trigger for the nominative form of the numeral, but not the genitive).

Since the North Sámi noun spells out to the right of the demonstrative and numeral (as well as the adjective, not shown), @ in the EP-span for the noun is relatively low, for example in *n*. I will assume that w is there as well. The case suffix is the exponent of a K head higher up, which bears no @ and therefore spells out as part of the noun word. We can assume that case markers on N in North Sámi are portmanteaux for D-K, since North Sámi doesn't generally use determiners (assuming D to be needed semantically).

(8)



Adnominal modifiers show partial case agreement, as seen in (6)–(7), so they must have a probe for case at the top of their extended projections. I designate the head with the case probe on the demonstrative D/D (a head which turns a demonstrative base into a D-adjunct, something which can be adjoined to a DP), and the head with the case probe on the

numeral Cl/Cl (a head which turns a numeral base into something which can be adjoined to a ClassifierP). The demonstrative and the numeral form prosodic words, so somewhere in their extended projections they contain a head with w, and they are not incorporated, so they also contain a head with @. I place both properties on the lower of the two heads depicted.

Now there are three @'s in this tree, and three words: Dem-D/D spells out as *man*, Num-Cl/Cl spells out as *golmma*, and N-n-Cl-D-K spells out as *dállui*. This means that the best exponent for D/D available is the underspecified *-n* which also spells out genitive, accusative, and locative on adnominal pronouns (including demonstratives), and the best exponent for Cl/Cl is zero, which also spells out genitive and locative on adnominal numerals. This in turn means that there is no more fully specified exponent for D/D or for Cl/Cl (following Caha 2009, we can take illative to be a composite of genitive, locative, and illative features, and assuming the subset principle of Halle and Marantz 1993, an exponent specified for genitive can spell out a node with all three, in the absence of a more fully specified exponent).

Now consider what happens if the noun is elided, as in (7b). Refer again to the tree in (8) but imagine that the nP is elided, with the result that the numeral 'three' shows illative case, *golbmii* instead of (genitive) *golmma*.

The omission of the nP, which contains the only @ in its EP-span, has the effect of stranding the span Cl-D-K. Under the narrow definition of syntactic word in (3) used up to this point, the span Cl-D-K is part of no word.

The distribution of case marking in North Sámi shows that the illative case exponent *-ii* plus a strong grade trigger, which can be taken to spell out Cl-D-K with illative features, ends up on the rightmost element remaining in the noun phrase, in this case the numeral. This suggests that an @-span can 'restructure', in the sense that when a stranded span with no @ has no complement, it can be spelled out as if it were the projection of its lowest specifier. In that case, the span is Num-Cl/Cl-Cl-D-K. Even though there is no specifically illative exponent for Cl/Cl, there is one for K, the one that ordinarily appears on nouns. Since the exponent of Cl/Cl was zero, this can simply be concatenated. Alternatively, we can assume that Cl/Cl and Cl are indistinguishable in category for the purposes of lexical insertion. In that case, a lexicalizer for Cl will also count as a lexicalizer for Cl/Cl, and no additional exponents will be needed for the adjunction head.

This gives good results in the case of the demonstrative, as seen in (7c). In the absence of a numeral, the demonstrative can bear illative. The restructured span is Dem-D/D-D-K. The illative exponent does not appear outside the underspecified *-n* (**ma-n-sa*). Instead, it appears adjacent to the demonstrative stem, and furthermore it has the form *-sa* which is used only with pronouns.

This suggests a possible revision of the definition of R_C in (1b). Instead of h_{n-1} being the head of the complement of h_n, h_{n-1} could be the head of the primary dependent of h_n, where primary dependent is a complement if present, adjunct or specifier otherwise. More cases need to be studied before it can be determined whether this change is warranted.

Further support for this analysis is provided by the illative plural, where case and number agreement are observed on adnominal modifiers, as seen in (9).

(9) ma-idda dálu-i-de
which-PL.ILL farm-PL-ILL
'to which farms'

The divergence between the singular and the plural suggests an account centered on properties of the exponents. This account promises to extend to other cases of edge inflection such as the one from Kuuk Thaayorre discussed in Inkelas 2017.

6. Conclusion

I have argued that a span is a thing, and sketched a span-based theory of words. The span-based theory of words refers to spans in several contexts. If just one of those is correct, then it vindicates the eponymous claim of the paper. In other words, any of several analyses here could be wrong, but if any one of my accounts of portmanteaux, of words, or of contextual allomorphy is right, then a span is nevertheless a thing.

References

Abels, Klaus, and Peter Muriungi. 2008. The focus marker in Kîîtharaka: Syntax and semantics. *Lingua* 118:687–731.

Adger, David. 2003. *Core syntax: A minimalist approach*. Oxford: Oxford University Press.

Adger, David. 2014. Variability and grammatical architecture. In *Linguistic variation in the minimalist framework*, ed. by M. Carme Picallo, 179–196. Oxford: Oxford University Press.

Adger, David, Daniel Harbour, and Laurel J. Watkins. 2009. *Mirrors and microparameters: Phrase structure beyond free word order*. Cambridge, UK: Cambridge University Press.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In *The morphology and phonology of exponence*, ed. by Jochen Trommer, 8–83. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 2017. Stratal phonology. In *The Routledge handbook of phonological theory*, ed. by S.J. Hannahs and Anna R. K. Bosch, 100–134. London: Routledge.

Bonet, Eulàlia, and Daniel Harbour. 2012. Contextual allomorphy. In *The morphology and phonology of exponence*, ed. by Jochen Trommer, 195–235. Oxford: Oxford University Press.

Borer, Hagit. 2005. *In name only (Structuring sense, vol. i)*. Oxford: Oxford University Press.

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

Bye, Patrik. 2007. Allomorphy—selection, not optimization. In *Freedom of analysis?*, ed. by Sylvia Blaho, Patrik Bye, and Martin Krämer, 63–91. Berlin: Mouton de Gruyter.

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

Caha, Pavel. 2009. The nanosyntax of case. Doctoral dissertation, University of Tromsø.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond: The cartography of syntactic structures, vol. 3*, ed. by Adriana Belletti, 104–131. Oxford: Oxford University Press.

Compton, Richard, and Christine Pittman. 2010. Word-formation by phase in Inuit. *Lingua* 120:2167–2192.

Dixon, R. M. W. 1972. *The Dyirbal language of North Queensland*. Cambridge, UK: Cambridge University Press.

Grimshaw, Jane. 2005. Extended projection. In *Words and structure*, ed. by Jane Grimshaw, 1–73. Stanford, Ca.: CSLI. Revised version of 1991 ms.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.

Hannahs, S. J., and Maggie Tallerman. 2006. At the interface: Selection of the Welsh definite article. *Linguistics* 44:781–816.

Harðarson, Gísli Rúnar. 2017. Cycling through grammar: On compounds, noun phrases and domains. Doctoral dissertation, University of Connecticut, Storrs.

Inkelas, Sharon. 2017. The directionality and locality of allomorphic conditioning in Optimal Construction Morphology. In *The morphosyntax-phonology connection: Locality and directionality at the interface*, ed. by Vera Gribanova and Stephanie S. Shih, 325–359. Oxford: Oxford University Press.

Julien, Marit. 2005. *Nominal phrases from a Scandinavian perspective*. Amsterdam: John Benjamins.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In *Linguistics in the morning calm*, ed. by I. S. Yang, 1–91. Seoul: The Linguistic Society of Korea and Hanshin.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, ed. by Alexis Dimitriadis and Laura Siegel, 201–225. Philadelphia: University of Pennsylvania.

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

Marvin, Tatjana. 2002. Topics in the stress and syntax of words. Doctoral dissertation, MIT, Cambridge, MA.

Merchant, Jason. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46:273–303.

Michael, Lev. 2014. The Nanti reality status system: Implications for the typological validity of the realis/irrealis contrast. *Linguistic Typology* 18:251–288.

Newell, Heather. 2008. Aspects of the morphology and phonology of phases. Doctoral dissertation, McGill University, Montréal.

Nickel, Klaus Peter. 1994. *Samisk grammatikk*. Karasjok: Davvi Girji, second edition.

Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70:491–538.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

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

Svenonius, Peter. 2004. On the edge. In *Peripheries: Syntactic edges and their effects*, ed. by David Adger, Cécile de Cat, and George Tsoulas, 261–287. Dordrecht: Kluwer.

Svenonius, Peter. 2008a. Paradigm generation and Northern Sámi stems. In *Inflectional identity*, ed. by Asaf Bachrach and Andrew Nevins, 73–100. Oxford: Oxford University Press.

Svenonius, Peter. 2008b. The position of adjectives and other phrasal modifiers in the decomposition of DP. In *Adjectives and adverbs: Syntax, semantics, and discourse*, ed. by Louise McNally and Chris Kennedy, 16–42. Oxford: Oxford University Press.

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

Svenonius, Peter. 2016a. Complex predicates as complementation structures. In *Approaches to complex predicates*, ed. by Léa Nash and Pollet Samvelian, 212–247. Leiden: Koninklijke Brill.

Svenonius, Peter. 2016b. Spans and words. In *Morphological metatheory*, ed. by Heidi Harley and Daniel Siddiqi, 199–220. Amsterdam: John Benjamins.

Svenonius, Peter. 2017a. Declension class and the Norwegian definite suffix. In *The morphosyntax-phonology connection: Locality and directionality at the interface*, ed. by Vera Gribanova and Stephanie S. Shih, 325–359. Oxford: Oxford University Press.

Svenonius, Peter. 2017b. V to C and late syntactic word formation. Ms., University of Tromsø. Paper presented in September 2017 at the *Comparative Germanic Syntax Workshop* at NTNU in Trondheim.

Svenonius, Peter. 2018. Delimiting the syntactic word. Ms., University of Tromsø. Paper presented in March 2018 at the University of California at Santa Cruz. Available at http://ling.auf.net/lingbuzz/003934.

Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral dissertation, MIT, Cambridge, MA.

Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 251–282. Cambridge, MA: MIT Press.

Vikner, Sten. 2005. Immobile complex verbs in Germanic. *Journal of Comparative Germanic Linguistics* 8:83–115.

Williams, Edwin. 2014. Derivational prefixes are projective, not realizational. In *Functional structure from top to toe: A festschrift for Tarald Taraldsen*, ed. by Peter Svenonius, 290–315. Oxford: Oxford University Press.

Peter Svenonius
peter.svenonius@uit.no