Spanning

Peter Svenonius
*CASTL, University of Tromsø**

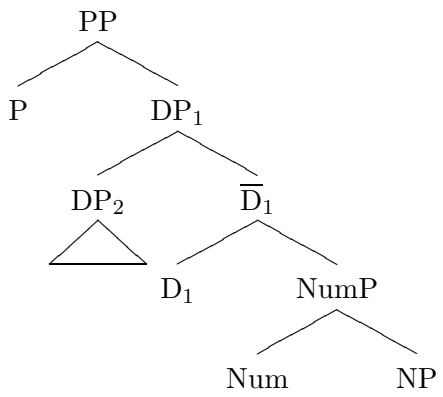
April 25, 2012

1 Introduction

I argue that the operation of Spell-out recognizes SPANS, defined in terms of a complement sequence of heads, normally in a single extended projection.[1] An extended projection (Grimshaw 2005) consists of a lexical head such as N or V and its associated functional projections (such as D, for N, and T, for V). A functional sequence is a partial (or total) ordering of a set of syntactic categories, for example if the nominal functional sequence is P↪D↪Num↪N, then PP, DP, and NumP are possible extended projections of N.[2] In an extended projection, each element in a functional sequence takes some lower element in the same sequence as a complement. A span is a subpart of the complement line in an extended projection.

For example, in the following tree (for a PP containing a noun phrase DP_1 with a prenominal possessor DP_2), P–D_1 is a span, and so is D_1–Num–N.

(1)



To be more precise, the exhaustive set of nontrivial spans in the main projection line in (1) is as follows (there will be a distinct set of spans for DP_2, nontrivial if D_2 has a complement):

(2)

(i)	P–D_1				
(ii)	P–D_1–Num	(iv)	D_1–Num		
(iii)	P–D_1–Num–N	(v)	D_1–Num–N	(vi)	Num–N

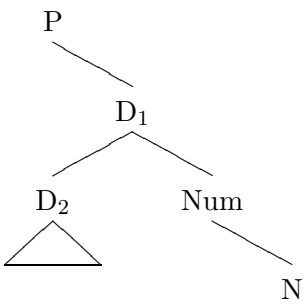
*Thanks to Tarald Taraldsen, Marina Pantcheva, Michal Starke, Patrik Bye, Heidi Harley, and two anonymous reviewers for discussion and comments on a previous draft.

[1]The term span is sometimes used for portmanteaux (in Hockett's 1947 sense, meaning a single morpheme which corresponds to two or more morphological 'slots') in the context of templatic descriptions of morphology, for example by Melnar (2004, 17). It is also used by Williams (2003, 214) for head-complement sequences within an extended projection, and is specifically used in the context of spell-out.

[2]For the purposes of this paper it does not matter whether the order is total or partial, or whether a given extended projection can be missing elements in the functional sequence.

This is more transparent if the tree is represented mirror-theoretically, eliminating redundant phrase labels and representing complementation with right-sloping lines (Brody 2000a,b).

(3)



A head is a trivial span. I suggest that morphological exponents are always associated with spans, trivial or nontrivial (see also Bye and Svenonius in press, where this assumption is explicit but not specifically argued for). A single morphological exponent (morpheme, for short) cannot spell out two heads (cannot "span" two heads) unless those heads are in a complement relation with each other. Thus, a single morpheme cannot spell out a head in an extended projection together with all or part of a specifier, nor can a single morpheme spell out a head in an extended projection together with all or part of an adjunct.[3]

The span is recognized indirectly in the theory of head-movement, since a head can normally only move to the head which selects it (Travis 1984; Baker 1988). Thus, spans can be thought of, roughly, as potential complex heads. However, the identity is not perfect; clitics can join in head-clusters even when they are not part of spans (see Roberts 2010 for a recent proposal and references). Furthermore, spans cannot be derived from head movement, as spans can be shown to be relevant to Spell-out even when no head-movement takes place (I discuss this distinction between spanning and head movement in section 4). in Brody's Mirror Theory, head-movement is a matter of where in a span a word linearizes; Brody (2000a,b) argues for example that a V-T span spells out parametrically in V (English) or T (French) (see also Adger et al. 2009 for an application of this to head-final structures).

In this paper I show how spanning accounts for the distribution of the French portmanteau preposition-determiner forms such as *du* and *au*. First, I lay out some background assumptions regarding the two-step procedure of Spell-out, returning to French preposition-determiners in section 3.

2 Spell-out in two steps

Spell-out maps syntactic structures onto representations which are the input to the phonological component (Chomsky 1993). The strict separation of phonology and syntax (Zwicky and Pullum 1986; Pullum and Zwicky 1988) motivates a two-step approach to Spell-out, where the first step (L-Match) associates syntactic structures with the syntactic features in lexical entries, and the second step (Insert) arranges the phonological content of those lexical entries into a form which is legible to phonology. This can be modeled in a version of Distributed Morphology

[3]For the purposes of this paper a span can be assumed to be confined to a single extended projection, without embedding of extended projections. However, there are cases in which a single morpheme appears to spell out a head such as V along with a head or heads in the extended projection of its complement (e.g. Gruber 1965; Hale and Keyser 2002; Son and Svenonius 2008). I will not discuss those cases in this paper but will assume for the time being that c-selection essentially turns a selected complement into part of the extended projection, at least for the purposes of lexical insertion (see Baltin 1989; Svenonius 1994 on the relationship of c-selection and the complement configuration).

(DM), and I will adopt the specific implementation detailed in Bye and Svenonius (2010, in press), including the OT model of the phonological component.

The model assumes late insertion, meaning that the list of exponents in the lexicon is accessed only after syntactic operations are complete in the relevant domain. The model also assumes cyclic Spell-out, meaning that the domain relevant for lexical access is the phase, a syntactic subconstituent of the utterance (Chomsky 2001).

The French determiner system distinguishes masculine from feminine (which I will represent using ±F), definite from indefinite (±DEF), and singular from plural (±PL). However, not all feature combinations are distinguished at once. Applying the algorithm for feature assignment developed in Adger (2006), we arrive at the following lexical entries.

(4)

a.	*un*	$\langle \mathrm{D}_{[-\mathrm{DEF},-\mathrm{F},-\mathrm{PL}]} \rangle$	⇔ /œ̃/
b.	*une*	$\langle \mathrm{D}_{[-\mathrm{DEF},+\mathrm{F},-\mathrm{PL}]} \rangle$	⇔ /yn/
c.	*le*	$\langle \mathrm{D}_{[+\mathrm{DEF},-\mathrm{PL}]} \rangle$	⇔ /lə/
d.	*la*	$\langle \mathrm{D}_{[+\mathrm{DEF},+\mathrm{F},-\mathrm{PL}]} \rangle$	⇔ /la/
e.	*les*	$\langle \mathrm{D}_{[+\mathrm{DEF},+\mathrm{PL}]} \rangle$	⇔ /le(z)/
f.	*des*	$\langle \mathrm{D}_{[-\mathrm{DEF},+\mathrm{PL}]} \rangle$	⇔ /de(z)/

These entries reflect the assumption (following e.g. Tranel 1996) that liaison with *les* and *des* (e.g. *les enfants* /lezãfã/ 'the children') reflects the presence of a structurally 'deficient' /z/, written /(z)/ in the lexical entry, which is deleted if it cannot be parsed as an onset (as in *les garçons* /legarsõ/ 'the boys').[4] Similarly, *le* and the prevocalic alternate *l'* are assumed to correspond to a single lexical entry with a deficient vowel /ə/, one lacking a root node (Zoll 1993) or a mora (Tranel 1996). Before another vowel, it elides, but before a consonant, it surfaces as schwa. No similarly regular phonological alternation could reduce feminine *la* to *l'*. Since *l'* appears with both masculine and feminine nouns, the entry for *le* is underspecified for gender, just like *les* and *des*. Its interaction with phonology is detailed below.

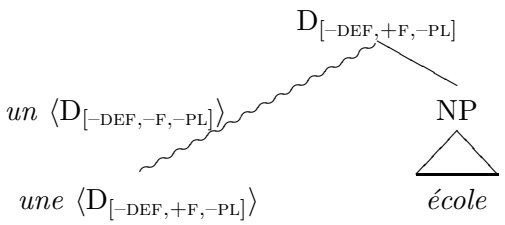
2.1 L-Match

The first step in each cycle of Spell-out is called L-Match (L for lexical). In each phase, the syntactic structure must be lexicalized, that is, linked to lexical entries. Lexical entries pair syntactic and phonological information, but L-Match only operates on features on the syntactic side of the pairing.

In (5), the operation L-Match is illustrated for a form like *une école* 'a school (F)' with squiggly lines representing lexical association to the syntactic tree, assuming the lexical entries in (5). The additional lexical entries contain features which do not match the tree, hence remain unassociated. The noun is assumed to have spelled out on a previous cycle (following Marantz 2007, who argues that all nouns contain a phase head *n*; see Marvin 2002; Newell 2008 on cyclic spell-out of words).

[4]The phonological analysis works the same whether or not *les* and *des* are bimorphemic and /(z)/ is a plural morpheme, a detail which I will set aside here.

(5)

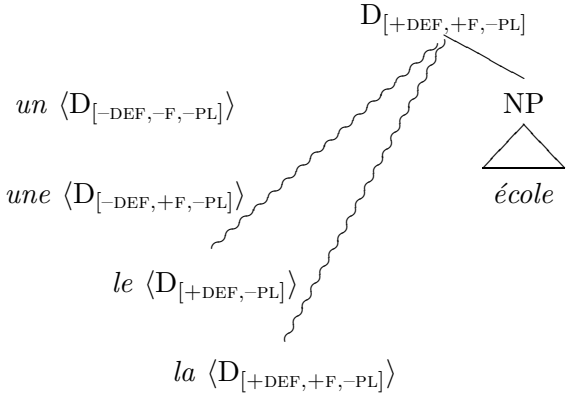


le $\langle D_{[+DEF,-PL]} \rangle$

la $\langle D_{[+DEF,+F,-PL]} \rangle$

Still following Adger (2006), nothing guarantees that there is a one-to-one function from syntactic trees to strings of exponents. L-Match can associate more than one exponent with a node in a tree. For example, as noted above, the definite article *le* is not marked for gender. In a definite singular noun phrase with a feminine noun, the set of associations illustrated in (6) arises.

(6)



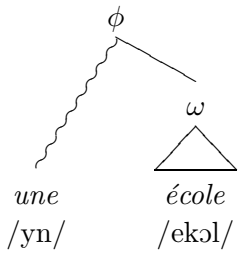
At this point I depart from a standard assumption in DM. In DM, conflicts like the one in (6) are normally assumed to be resolved by reference to properties of the syntactic features involved. In most such cases, the Elsewhere Principle systematically chooses the more fully specified alternative, blocking the other option from occurring (Halle and Marantz 1993). This kind of blocking can be straightforwardly computed only when the features on one lexical entry are a subset of the features on another. In the case in (6), the Elsewhere Principle would force the choice of *la*, and the association link to *le* would not be retained. If this model is to be reconciled with the assumption that the Elsewhere Principle operates in L-Match, then I would have to reconfigure the lexical entries here so that neither was a subset of the other, for example by positing an additional feature on *le*. However, I will instead assume that L-Match tolerates indeterminacy, and that the Elsewhere Principle is not invoked at this level (essentially following Adger 2006).[5]

[5]Adger (2006) does not discuss phonologically conditioned allomorphy, but uses underspecification of this kind for sociolinguistic variability. What I am suggesting here is that as far as the syntax is concerned, the alternation between *la* and *l'* with feminine nouns can be treated the same way, letting phonology rather than sociolinguistic factors sort out the distribution.

2.2 Insert

Once the exponents have been matched systematically in this way, Spell-out proceeds to a second step, called Insert in Bye and Svenonius (2010, in press), in which the syntactic features are no longer available and only the phonological qualities of the matched exponents are relevant. At this point, even the syntactic nodes themselves become irrelevant, as do their dependent syntactic features, as they are mapped onto phonological constituents (here, ω for phonological word and ϕ for phonological phrase, cf. e.g. Selkirk (2011); the details of whether ϕ is recursive or n-ary branching and so on are irrelevant for present purposes).

(7)



A linearization algorithm determines the order in which the phonological units are concatenated, on the basis of the nature of the syntactic dependencies (in the diagrams, squiggly lines are lexical dependencies, and straight lines are structural dependency relations — right-sloping for complements, and left-sloping for specifiers). For unincorporated complements, dominance maps to precedence: if α lexicalizes A (i.e. is a lexical dependent of A), then α precedes any material lexicalizing any complement of A. Specifiers do not figure into the examples in this paper, but I assume that the linearization algorithm has the effect that specifiers precede heads (as in Kayne 1994). I also follow Brody (2000a,b) in assuming that morphologically incorporated complements precede heads, as a principle of linearization. The temporal precedence relation is represented in the tree in (7) by left-right order of the phonological elements. This is the input to the phonological component (which may realign prosodic constituency, for example making the /n/ of /yn/ into the onset of the following syllable, hence realigning the phonological word boundary).

For L-Match, all features manipulated by syntax are relevant (including gender, which is copied by Agree), but L-Match is blind to phonological features. L-Match cannot resolve cases of contextual allomorphy, where the phonological environment determines which allomorph is selected. A well-known example is the English indefinite article, where *a* /ə/ is used before consonants and *an* /ən/ before vowels. In such cases, the output of L-Match can be assumed to include two options, resolved later in the phonology. Thus, the two options {*a, an*} must be indistinguishable to the syntax, and so the single lexical entry for the indefinite article can simply include both options.

(8) *a, an* $\langle \mathrm{D}_{[-\mathrm{DEF}]} \rangle \Leftrightarrow$ {/ə/, /ən/}

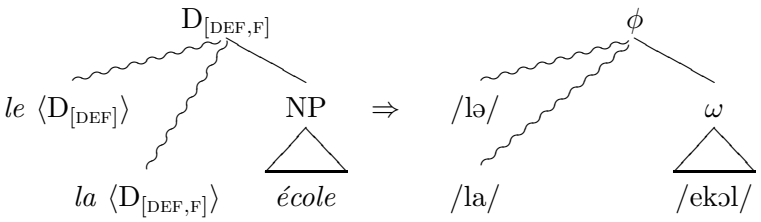
Either the marked form has a phonological condition associated with it (Hayes 1990), or else both forms are entered into Gen and evaluated together by the phrasal phonology (Mascaró 1996).

Recall from the above that the alternation in masculine nouns between *le* and *l'* (as in *le garçon* 'the boy,' *l'enfant* 'the child') is not a case of contextual allomorphy, but of pure phonology; the 'defective' segment /ə/ is deleted unless required to support syllable structure. But the choice between *la* and *l'* with feminine nouns (as in *la profession* 'the profession,' *l'usine* 'the factory') is a case of contextual allomorphy just like the choice between *a* and *an*

in English.

The L-Match tree from (6) is repeated here in (9). It is converted in Insert to the structure on the right.

(9)



Lexicalizations of ϕ are linearized before material under ω, as expected, but the multiple exponents associated with ϕ are not linearized with respect to each other. This provides Insert with two options for lexicalizing the ϕ node, *la* and *le*; Gen applies to both of these options to generate multiple candidates for pronunciation. Candidates in which *le* appears without its defective vowel violate a low-ranked constraint against deleting defective segments, which I call PARSE-V here, following van Oostendorp (2000) (deleting the /a/ in *la*, on the other hand, would additionally violate a higher-ranked PARSE constraint against deleting fully specified segments, following Tranel (1996), not shown here).

(10)

/{lə, la} ekɔl/	ONS	PARSE-V	*ə
a. ☞ le.kɔl		*	
b. lə.e.kɔl	*!		*
c. la.e.kɔl	*!		

I have also included a relatively low-ranked constraint against realizing the defective vowel as a schwa (Tranel 1996 posits a constraint AIF 'avoid integrating floaters,' and van Oostendorp 2000 uses '*[–cons]' to punish vocalic roots; I simply call whatever constraint is responsible *ə, and place it lower than PARSE-V on the basis of forms like *prenez-le* /prənelə/, 'take it'). This set of rankings causes *la* to be preferred over *le* (the ONSET violations in (11) are for the syllables consisting only of /l/).

(11)

/{lə, la} tabl/	ONS	PARSE-V	*ə
a. l.tab.l	**!	*	
b. lə.tab.l	*		*!
c. ☞ la.tab.l	*		

There are several cases like this in French where adnominal elements have two contextual allomorphs, one prevocalic and one preconsonantal. Tranel (1996) and Mascaró (1996) discuss what the latter calls '*belle*-allomorphy,' based on alternations like the following.

(12)

	Masculine nouns		Feminine nouns	
	C-initial	V-initial	C-initial	V-initial
'beautiful'	*beau curé*	*bel abbé*	*belle table*	*belle arme*
	'parson'	'abbot'	'table'	'weapon'

A closed class of adjectives in French show this pattern (*noveau∼nouvelle* 'new,' *vieux∼vieille* 'old,' *fou∼folle* 'crazy,' etc.). As Mascaró notes, *bel* /bɛl/ cannot be derived from the same base as *beau* /bo/ by any rules motivated in French phonology, so the alternatives must be lexically listed suppletive forms. However, feminine *belle* and masculine prevocalic *bel* are phonologically identical, /bɛl/, and so can be treated as orthographic variants (which I will represent *bel(le)*). The learner, then, is faced with an alternation between /bo/ and /bɛl/ with masculine nouns, in syntactically identical contexts. Thus, according to the algorithm adopted from Adger (2006), the form /bɛl/ is underspecified for gender, while *beau* /bo/ is specified as [−F].

This ensures that *beau* is prevented in L-Match from appearing with feminine nouns, and so only *bel(le)* can surface there. But with masculine nouns, both alternants are associated by L-Match, just as happened with feminine nouns and the definite article, with the version being chosen which gives the best syllable structure. This is formally represented by Mascaró as follows for *bel ami* 'beautiful friend (M)' and *beau mari* 'beautiful husband (M),' omitting several constraints which do not come into play here.

(13)

/{bɛl, bo} ami/	Ons	No-Coda
a. ☞ bɛ.la.mi		
b. bo.a.mi	*!	

(14)

/{bɛl, bo} mari/	Ons	No-Coda
a. bɛl.ma.ri		*!
b. ☞ bo.ma.ri		

Zwicky (1987) analyzes this same data set, without using underspecification or a two-step lexical insertion procedure. Instead, he posits a 'Rule of referral' which stipulates that the feminine form is substituted for the masculine in prevocalic cases (his (53), p. 223, reads, "[MINOR, +DEF, MASC, SG] in morphosyntactic structure is referred to the corresponding [FEM] when a V-initial word follows," where MINOR is a diacritic to mark the special function words to which the rule applies). The rule invokes both syntactic and phonological information in a single statement, something which I am proposing to do without. Furthermore, it cannot be extended to the definite article, or to *mon* 'my' which also shows a "masculine" form being substituted for the feminine (e.g. *mon arme* 'my weapon (F),' in Tranel's example).

(15)

	Masculine nouns		Feminine nouns	
	C-initial	V-initial	C-initial	V-initial
'my'	*mon curé*	*mon abbé*	*ma table*	*mon arme*
	'parson'	'abbot'	'table'	'weapon'

Tranel (1996) proposes an OT analysis in which agreement in gender can be ranked among phonological constraints, which again violates the separation of phonology and syntax. The account of Perlmutter (1998) similarly mixes phonological and morphological constraints. Furthermore, these proposals overgenerate unless the constraints are arbitrarily associated with individual lexical items, as the substitution only occurs for a fixed set of adnominal elements. Mascaró (1996) points out that it does not apply, for example to *favori* 'favorite M' and *favorite* 'favorite F,' e.g. it is *favori ami* 'favorite friend (M),' not the phonologically superior **favorite*

ami.[6] The account proposed here does entirely without indexed constraints (following Bye and Svenonius 2010, in press, and references there).

The analysis proposed here holds that a learner posits underspecified entries for forms like *le*, *les*, *mon*, *bel(le)*, and so on, because they are observed to appear with both genders of noun. Forms like *la*, *ma*, *beau*, and so on are specified for gender because they consistently appear with nouns of a particular gender. In cases where L-Match finds two suitable exponents, Insert may distinguish them on the basis of phonology (otherwise, they will presumably appear as variants, as in the cases discussed in Adger 2006).

This analysis is therefore not consistent with common practice in DM, where underspecification is used extensively and the Elsewhere Principle is assumed to adjudicate at what corresponds to the L-Match level here (Halle and Marantz 1993). I am thus assuming a more restricted role for underspecification and the Elsewhere Principle (it may apply to certain kinds of features, for example categorial features, or interpretable features, if the ϕ-features on D can be assumed to be uninterpretable manifestations of features interpretable elsewhere; space prevents me from pursuing this matter further here).

3 Preposition-determiner fusion

Having established how the two-step model of Spell-out works, I turn to portmanteaux.

It is relatively common for prepositions and determiners to interact morphologically (cf. e.g. Napoli and Nevis 1987 on Italian, Waldmüller 2008 on German). In a cyclic model of spell-out, the interactions suggest that they are in the same phase and are located in the same spell-out domain.[7] A well-known example is the French fusion of the masculine definite article with the prepositions *de* and *à*. This happens only if the determiner is not contracted with a vowel-initial noun phrase, as seen in the table below (from Zwicky 1987, 212).

(16)

	Feminine nouns		Masculine nouns	
	V-initial	C-initial	V-initial	C-initial
	l'école	*la maison*	*l'hôpital*	*le parc*
	/lekɔl/	/lamɛzɔ̃/	/lɔpital/	/ləpark/
à	*à l'école*	*à la maison*	*à l'hôpital*	*au parc*
	/alekɔl/	/alamɛzɔ̃/	/alɔpital/	/opark/
de	*de l'école*	*de la maison*	*de l'hôpital*	*du parc*
	/dəlekɔl/	/dəlamɛzɔ̃/	/dəlɔpital/	/dypark/

The pattern shows that the phonology of the noun phrase is relevant to the allomorphy of the preposition: if the noun phrase is vowel-initial, then the determiner is procliticized and the preposition is regular; if the noun phrase is not vowel-initial, then in the right context of morphological features (here, masculine singular; there is a parallel interaction in the plural,

[6] Mascaró (1996) proposes to treat *belle*-allomorphy in French much like the English {*a, an*}, simply listing {*beau, bel*} as alternate allomorphs of the masculine adjective, and leaving the phonological identity of *belle* and *bel* (and other such cases) to factors outside the grammar proper (e.g. historical factors). Adopting such a proposal would leave unaffected the main point of this squib concerning spans. See also Mascaró (2007); Bonet et al. (2007) for a development in which allomorphs are partially ordered by a preference relation, and adherence to this preference can be interranked with phonological constraints.

[7] Law (1998) suggests languages with fused P-D elements never allow P-stranding. If that is correct, then it suggests a syntactic prerequisite for P-stranding which is incompatible with P-D fusion. In my account, such a prerequisite could be DP having phase status independent of PP (e.g. in English, which allows P-stranding, but not in German, which has fused P-D elements). However, Law's generalization is based on a very small set of European languages, and so the correlation could be accidental.

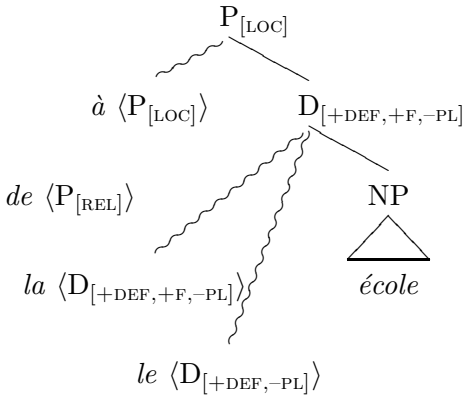
not illustrated here), a special portmanteau allomorph is chosen. Thus, allomorph selection for the adposition must happen after the phonological content of the noun phrase has become accessible.

The two prepositions involved can be given the following entries, letting "[REL]" and "[LOC]" stand in for whatever features distinguish the two.

(17) a. *de* $\langle \mathrm{P}_{[\mathrm{REL}]} \rangle \Leftrightarrow$ /də/
b. *à* $\langle \mathrm{P}_{[\mathrm{LOC}]} \rangle \Leftrightarrow$ /a/

For a form like *à l'école*, L-Match associates the entry in (17b) to a tree like that in (18).

(18) L-Match



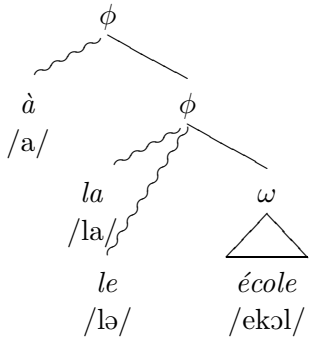
les $\langle \mathrm{D}_{[+\mathrm{DEF},+\mathrm{PL}]} \rangle$

...

Given the lexical items listed here, the only match for P is *à*, but D has two matches, the more specified *la* and the less specified *le*. The plural form *les* is not a match, on these assumptions, since it bears a plural feature. Nor is the preposition *de*, because it also has a feature, [REL], which is not matched in the syntactic tree.

Once the exponents have been matched syntactically in this way, spell-out proceeds to Insert, in which the phonological qualities of the matched exponents are relevant, but not the syntactic features.

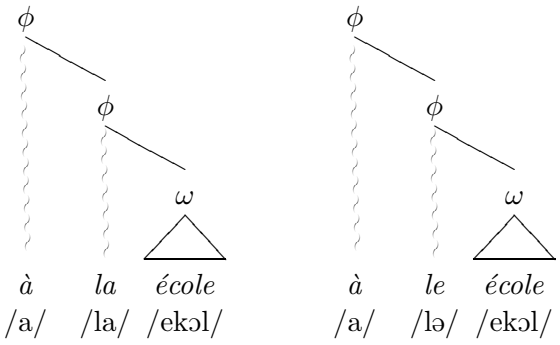
(19) Insert



As before, the linearization algorithm, which puts unincorporated higher heads before lower structure, linearizes /a/ to the left of both /la/ and /lə/, and all of those to the left of /ekɔl/,

but does not linearize /la/ with respect to /lə/. They depend on a single node, and only one exponent is needed to lexicalize that node, so they compete. Thus, we now compare /alaekɔl/ to /aləekɔl/ (along with the prosodic structure represented in (19)). This is easier to see if we redraw (19) as a set of trees in which each node is represented only by one exponent.

(20) Insert



Alternatively, we can represent the same information in the form of a tableau. Here, for simplicity of formatting, I omit the prosodic structure represented in the trees from the input line of the tableau.

(21)

/a {lə, la} ekɔl/	DEP	ONS	PARSE-V	*ə
a. ☞ a.le.kɔl		*	*	
b. ta.le.kɔl	*!		*	
c. a.lə.e.kɔl		**!		*
d. a.la.e.kɔl		**!		

Essentially the same factors adjudicate here as in the form for *l'école* seen in (10). I have added a DEP constraint against epenthesis in order to show a candidate, ((21)b), which does not violate ONSET.

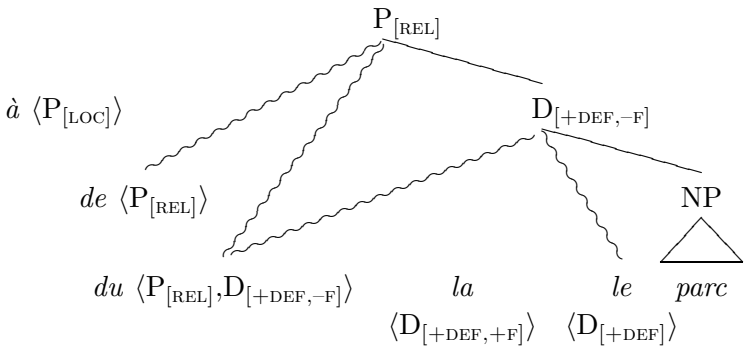
In addition to the regular D's and P's, there are the portmanteau forms, which associate to the span of P–D. This is specified in their lexical entries (compare the proposals of Zwicky 1987 and of Mascaró 1996).

(22) a. *au* $\langle P_{[LOC]}, D_{[+DEF,-F,-PL]} \rangle \Leftrightarrow$ /o/
b. *du* $\langle P_{[REL]}, D_{[+DEF,-F,-PL]} \rangle \Leftrightarrow$ /dy/

First, note that the new entry *au* does not falsely predict **au école*: because *au* is specified as masculine, it will never be associated to the tree with a feminine feature on D.

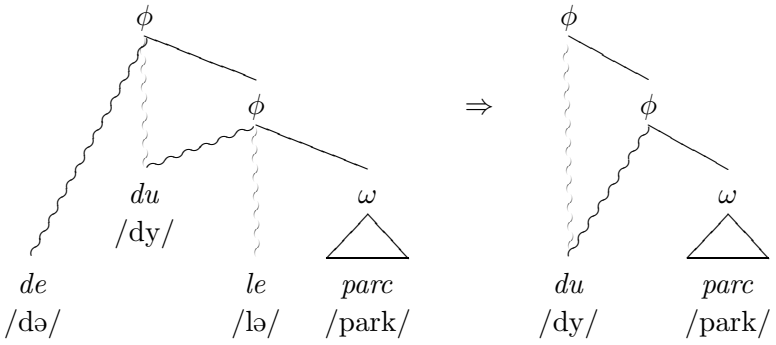
Now consider the spell-out of *du parc* (setting aside the ±PL features, for simplicity). Here, because the entry for *du* specifies both P and D features (it is a portmanteau), L-Match associates it to two nodes. At the same time, *de* is also a match, and so is *le*.

(23) L-Match



Moving to Insert, we discard the unassociated morphemes and consider only the phonological information in the remaining lexical entries. Although *de* can be linearized before *le*, and *le* before *parc*, and *du* before *parc*, there is no relative linearization of *de le* and *du*. They lexicalize the same nodes, as shown in the left-hand tree below, and thus they compete.

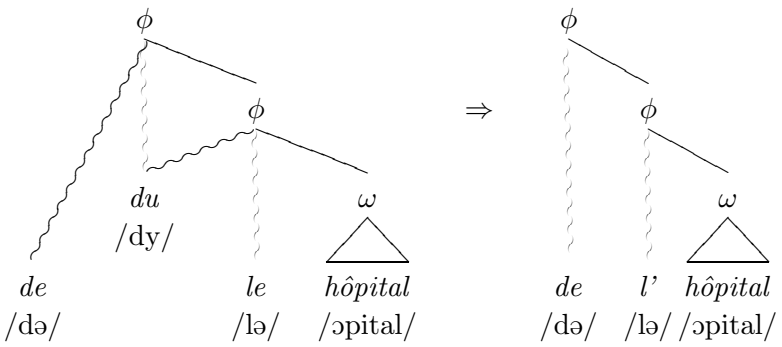
(24) Insert



The correct result is of course *du parc*, as indicated in the right-hand tree above. The question is why this is preferred over **de le parc*. There are many cases in which a portmanteau morph is observed to block an apparently equally fully specified sequence of two separate morphemes, and it has been proposed that there is a general principle favoring fewer exponents over more exponents, all else being equal (e.g. Muriungi 2009; Caha 2009; Siddiqi 2009; Taraldsen 2010). However, in the case at hand it can be seen (in (16)) that the principle is overridden by concerns of syllabic well-formedness, since vowel-initial masculine nouns like *hôpital* take the bimorphemic sequences *de l'*, *à l'* rather than *du*, *au*.[8]

(25) Insert

[8]As usual, the Insert trees represent intermediate stages in the phonological derivation. Syllable well-formedness constraints will ensure that the definite article *l'* is contained in the same phonological word as the vowel-initial noun *hôpital* in the final output of phonology, although in the righthand tree here it is represented as being outside the word.



Given the strict separation of syntax and phonology adopted here, this means that somehow, the phonological component must have a way of preferring fewer exponents over multiple exponents, if onset requirements are satisfied. I will assume that this is because phonology must concatenate sequences of exponents, and each instance of concatenation creates prosodic structure. In other words, the two ϕ nodes in (24) and (25) are distinct when associated to distinct exponents, but they are one and the same when only one exponent (*du*) is associated. A general constraint *Struct punishes candidates with more structure compared to those with less.

In the tableaux below, this is represented by assigning one violation mark under *Struct for each instance of concatenation. The constraint *Struct is crucially ranked below Onset; I place it above Parse-Vand *ə here for concreteness.

(26)

/{də lə, dy} park/	Ons	No-Coda	*Struct	Parse-V	*ə
a. ☞ dy.park		**	*		
b. də.lə.park		**	**!		**
c. dəl.park		***!	**	*	*

(27)

/{də lə, dy} ɔpital/	Ons	No-Coda	*Struct	Parse-V	*ə
a. dy.ɔ.pi.tal	*!	*	**		
b. də.lə.ɔ.pi.tal	*!	*	***		**
c. ☞ də.lɔ.pi.tal		***	***	*	*

4 Comparison with an alternative

Embick (2007, 2010) develops an alternative account within a more standard version of Distributed Morphology. In that framework, spanning is not possible, and lexical insertion is restricted to terminal nodes. However, certain operations may apply to fuse adjacent terminal nodes together, creating new, composite terminal nodes. One such operation is local dislocation, and Embick states two rules of local dislocation, specific to French, in (28).

(28) Embick (2010, 88)

a. Article Cliticization
D[def]$^\frown X \rightarrow$ [D[def][X]], X V-initial

b. P-D Affixation
P$^+{}^\frown$D[def]$^+ \rightarrow$ [P$^+$[D$^+$]]
where $^+$ is a diacritic for the particular terminals that are subject to this process

The first rule, Article Cliticization, states that if a noun phrase (X) is vowel-initial, the definite article (D[def]) will cliticize to it, bleeding the following rule. The second rule, P-D Affixation, states that certain specially-marked definite articles (namely, the masculine ones) form a constituent with an adjacent specially-marked P (namely, the ones that would otherwise spell out as *à* or *de*).

The first thing to note about these rules is that they are much more powerful than spanning, since they operate under adjacency; any two heads could be combined in this way, after syntactic movement, whereas spanning is restricted to spans (i.e. subsegments of a functional sequence). Therefore, the burden of proof should be on this version of DM to show that such powerful mechanisms are empirically necessary.

The second thing to note about these rules is that they mix syntactic and phonological information. If the formation of a terminal node is syntactic, then the reference to vowel-initial elements in (28a) is a violation of the well-motivated Principle of Phonology-Free Syntax (Zwicky and Pullum 1986; Pullum and Zwicky 1988), which states that syntactic rules do not make reference to phonology. If the formation of a constituent in (28a) is meant to be phonological, then the rule violates the converse of the same principle, formalized for example in Inkelas' (1989) Indirect Reference principle, which states that phonological rules can only refer to syntactic structure via prosodic constituency; (28a) refers to the syntactic features D and [def]. If there is a module of morphology which mixes syntax and phonology in the way that (28) does, then the empirical observations which motivate the Principle of Phonology-Free Syntax and the Indirect Reference principle are left unexplained.

I claim rather that the simple cliticization of the definite article is due to phonological properties of the exponent *le*, which has a defective vowel. Once the phonological properties of the element lexicalizing D have been accessed, a syntactic constituent cannot be formed out of P+D without opening up the possibility of phonology influencing syntax; therefore, the portmanteau status of P-D cannot be a syntactic fact, but is a fact about spell-out and the exponents it associates with syntactic structures, one which is handled here through spanning.

The problem for standard implementations of DM such as Embick's is essentially that if lexical insertion is to be applied to terminal nodes, then before lexical insertion occurs, there must be some independent organization of the syntax into terminal nodes. Here, items which are only visible through lexical insertion itself (namely the portmanteaux *au* and *du*) are the only clue that P and D must be reorganized into terminal nodes. Hence the diacritics '$^+$' in Embick's rules, which have no other motivation. All theories have to list irregular portmanteaux like *au* and *du*; on the spanning theory, that is all that need be done, no additional features need be posited.

5 Extending the account to *en*

In this section I briefly discuss a pattern concerning a class of proper place names which provides a few clues about the nature of L-Match and Insert and the lexical entries that it deals with. With certain proper place names (including many names of countries), the preposition *en* is ordinarily used, in lieu of *à* (examples from Zwicky 1987).[9]

[9]The exact distribution of this pattern is complex, with some place names showing different patterns. I will assume for the time being that the differences can be attributed to differences in the featural make-up of D, and restrict the analysis to the class discussed by Zwicky.

(29)

	Feminine place names		Masculine place names	
	V-initial	C-initial	V-initial	C-initial
	l'Amérique	*la France*	*l'Iran*	*le Canada*
	/lamerik/	/lafrɑ̃ːs/	/lirɑ̃/	/ləkanada/
LOC	*en Amerique*	*en France*	*en Iran*	*au Canada*
	/ɑ̃namerik/	/ɑ̃frɑ̃ːs/	/ɑ̃nirɑ̃/	/okanada/

The exception, as can be seen in the table, is masculine nouns beginning with consonants. This pattern can be straightforwardly described in the spanning model on the assumption that *en* is specified like *au* in being a P–D portmanteau, and in being specified as [LOC], but differing in the features present on the determiner. First, *en* is not restricted to masculines, as *au* is, and second, *en* is used in certain marked cases with 'defective' determiners. I will model this by specifying *en* with a distinct feature on D, [PRPR] (for 'proper'). The two entries are presented side-by-side for comparison.

(30) a. $\langle P_{[LOC]}, D_{[+DEF,-F]} \rangle \Leftrightarrow$ *au* /o/ (repeated from (22a))
b. $\langle P_{[LOC]}, D_{[PRPR]} \rangle \Leftrightarrow$ *en* /ɑ̃(n)/

The [PRPR] feature is intended to ensure that *en* is never a contender in the cases already discussed, since L-Match will not form associations with a lexical entry that is specified for features not appearing in the syntactic tree. By the same convention, *au* will not be considered for nonmasculine nouns, since it has a feature [–F]. Phonological competition in Insert will therefore only occur in the masculine, and only with the special D such as that occurring in this class of proper place names.[10]

In that narrow context, the empirical fact is that *en* beats both *au* and *à le* before vowels (*en Iran*), and *au* beats both *en* and *à le* before consonants (*au Canada*). This is consistent with the phonological assumptions already made here. First, consider the situation before vowels. The forms */oirɑ̃/ and */aləirɑ̃/ lose out because of ONSET; the form */alirɑ̃/ requires the moraless /ə/ to be deleted in order to avoid hiatus, violating PARSE-V, and also requires an extra application of *STRUCT compared to the form /ɑ̃nirɑ̃/. The latter requires a rootless consonant /(n)/ to receive a root node, violating a DEP-ROOT constraint, which must be ranked lower than one or the other or both of *STRUCT and some PARSE constraint.

Turning to the example with a consonant (*au Canada*), it violates ONSET once, but so do its closest competitors, and it introduces no other phonological complications, whereas both of the other forms require either deleting a (possibly rootless) consonant to avoid a coda (**en Canada*) or inserting a root or mora (**à le Canada*). Furthermore, the monomorphemic status of *au* gives it an additional edge over *à le*, because of there being one less *STRUCT violation.

6 Conclusion

I conclude that spans are recognized by the grammar: Lexical insertion targets spans, including trivial spans, of functional sequences (possibly including complement sequences derived through c-selection, not discussed here). This promises to shed new light on head movement, extended projections, and the status of such objects as morphemes, words, and lexical items. The significance of spans for grammar also suggests that the complement relation must be distinguished from the specifier relation; thus complements, at least within extended projections, are not

[10] Proper place names will also have to have the [+DEF] feature or else *au* would not be considered. The featural specifications used here are an oversimplification. For example, even the proper place names in (29) can appear with *à* in certain contexts (Zwicky mentions *penser à*, 'think about,' and toasts). These involve additional features which can be expressed by *la/le* but not *en*, which only spells out certain featurally restricted D.

simply the first-merged constituents, but have a distinct status. This has implications for the theories of selection and projection.

References

Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42 3: 503–530.

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

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

Baltin, Mark R. 1989. Heads and projections. In *Alternative Conceptions of Phrase Structure*, edited by Mark R. Baltin and Anthony S. Kroch, pp. 1–16. University of Chicago Press, Chicago.

Bonet, Eulàlia, Maria-Rosa Lloret, and Joan Mascarò. 2007. Allomorph selection and lexical preferences: Two case studies. *Lingua* 117: 903–927.

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*, edited by Peter Svenonius, pp. 27–43. John Benjamins, Amsterdam.

Bye, Patrik and Peter Svenonius. 2010. Exponence, phonology, and non-concatenative morphology. Ms. CASTL, University of Tromsø; available at ling.auf.net/lingbuzz/001099.

Bye, Patrik and Peter Svenonius. in press. Non-concatenative morphology as epiphenomenon. In *The Morphology and Phonology of Exponence: The State of the Art*, edited by Jochen Trommer, pp. 427–495. Oxford University Press, Oxford.

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

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 1–52. MIT Press, Cambridge, Ma.

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

Embick, David. 2007. Linearization and local dislocation: Derivational mechanics and interactions. *Linguistic Analysis* 33 3-4: 2–35.

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

Grimshaw, Jane. 2005. Extended projection. In *Words and Structure*, edited by Jane Grimshaw. CSLI, Stanford, Ca. Final version of 1991 ms.

Gruber, Jeffrey. 1965. *Studies in Lexical Relations*. Ph.D. thesis, Massachusetts Institute of Technology, Department of Modern Languages.

Hale, Ken and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. No. 39 in Linguistic Inquiry Monographs. MIT Press, Cambridge, Ma.

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*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 111–176. MIT Press, Cambridge, Ma.

Hayes, Bruce. 1990. Precompiled phrasal phonology. In *The Phonology-Syntax Connection*, edited by Sharon Inkelas and Draga Zec, pp. 85–108. University of Chicago Press, Chicago.

Hockett, Charles F. 1947. Problems of morphemic analysis. *Language* 23 4: 321–343.

Inkelas, Sharon. 1989. *Prosodic Constituency in the Lexicon*. Ph.D. thesis, Stanford, Palo Alto, Ca.

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

Law, Paul. 1998. A unified analysis of P-stranding in Romance and Germanic. In *Proceedings of the North East Linguistic Society*, edited by Pius N. Tamanji and Kiyomi Kusumoto, vol. 28, pp. 219–234. GLSA, Amherst, Ma.

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

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

Mascaró, Joan. 1996. External allomorphy and contractions in Romance. *Probus* 8 2: 181–206.

Mascaró, Joan. 2007. External allomorphy and lexical representation. *Linguistic Inquiry* 38 4: 715–735.

Melnar, Lynette R. 2004. *Caddo Verb Morphology*. University of Nebraska Press, Lincoln, Ne.

Muriungi, Peter Kinyua. 2009. The union spell-out mechanism. In *Tromsø Working Papers on Language and Linguistics: Nordlyd 36.1,* Special issue on Nanosyntax, edited by Peter Svenonius, Gillian Ramchand, Michal Starke, and Knut Tarald Taraldsen, pp. 191–205. University of Tromsø, Tromsø. Available at http://www.ub.uit.no/baser/nordlyd/.

Napoli, Donna Jo and Joel Nevis. 1987. Inflected prepositions in Italian. *Phonology* 4 1: 195–209.

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

van Oostendorp, Marc. 2000. *Phonological Projection: A Theory of Feature Content and Prosodic Structure*. Mouton de Gruyter, Berlin.

Perlmutter, David M. 1998. Interfaces: Explanation of allomorphy and the architecture of grammars. In *Morphology and its Relation to Phonology and Syntax*, edited by Steven G. Lapointe, Diane K. Brentari, and Patrick M. Farrell, pp. 307–338. CSLI, Stanford, Ca.

Pullum, Geoffrey K. and Arnold Zwicky. 1988. The syntax-phonology interface. In *Linguistics: The Cambridge Survey*, edited by Frederick J. Newmeyer, vol. 1, pp. 255–280. Cambridge University Press, Cambridge.

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

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In *The Handbook of Phonological Theory*, edited by John Goldsmith, Jason Riggle, and Alan Yu. Blackwell, Oxford, 2nd edn.

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

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*, edited by Natasha Abner and Jason Bishop, pp. 388–396. Cascadilla, Somerville, Ma.

Svenonius, Peter. 1994. C-selection as feature-checking. *Studia Linguistica* 48 2: 133–155.

Taraldsen, Knut Tarald. 2010. The nanosyntax of Nguni noun class prefixes and concords. *Lingua* 120: 1522–1548.

Tranel, Bernard. 1996. French liaison and elision revisited: A unified account within Optimality Theory. In *Aspects of Romance Linguistics*, edited by C. Parodi, C. Quicoli, M. Saltarelli, and M. L. Zubizarreta, pp. 433–455. Georgetown University Press, Washington, D.C.

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

Waldmüller, Estela. 2008. *Contracted Preposition-Determiner Forms in German: Semantics and Pragmatics*. Ph.D. thesis, Universitat Pompeu Fabra, Barcelona.

Williams, Edwin. 2003. *Representation Theory*. MIT Press, Cambridge, Ma.

Zoll, Cheryl. 1993. Ghost segments and optimality. In *The Proceedings of the West Coast Conference on Formal Linguistics*, edited by Erin Duncan, Donka Farkas, and Philip Spaelti, vol. 12, pp. 183–199. Cascadilla Press, Somerville, Ma.

Zwicky, Arnold. 1987. French prepositions: No peeking. *Phonology* 4: 211–227.

Zwicky, Arnold M. and Geoffrey K. Pullum. 1986. The Principle of Phonology-Free Syntax: Introductory remarks. *OSU Working Papers* 32: 63–91.