Lexical Insertion in two stages

Extended abstract

Peter Svenonius, CASTL, University of Tromsø

1   Introduction

I propose a new variant on the model of late lexical insertion, in which the insertion of phonological exponents occurs in cycles, from the bottom of the syntactic tree upward (as in Distributed Morphology (DM, Halle and Marantz 1993, Harley and Noyer 1999, Embick and Noyer 2007) and much other work). The size of the cycle is the 'phase' proposed by Chomsky (2001 inter alia), roughly the extended projection (in Grimshaw's 2005 sense) of a lexical category or of T or D. Within each cycle, individual exponents may associate with 'spans' of functional material larger than a single head, as argued in Nanosyntax (Caha 2009, Svenonius et al. 2009, Pantcheva 2011). Lexical insertion, drawing on Bye and Svenonius (2010; in press), takes place in two stages, a syntactic one which makes no reference to phonology (in accord with the *Principle of Phonology-Free Syntax* (PPFS) of Zwicky and Pullum 1986, Pullum and Zwicky 1988), followed by a phonological one which makes no reference to to syntax, in accordance with the *Indirect Reference Principle* (IRS, Inkelas 1989).

Thus, I argue against a central principle common to most work in Distributed Morphology (DM): DM assumes that lexical insertion occurs one terminal node at a time, and competition among allomorphs is resolved at each step, often by Kiparsky's (1973) Elsewhere Principle. This has been argued explicitly by Bobaljik (2000), Embick and Marantz (2008), for example. Bottom-up terminal node insertion predicts that there can be no selection of suppletive allomorphs which is sensitive to phonological properties 'outward' (in Carstairs' 1987 sense, i.e. further from the root). I show that Icelandic participle formation has exactly this property, supporting my claim that lexical insertion does not occur one node at a time, but one phase at a time. Terminal node insertion also predicts that portmanteau morphemes (i.e. exponents which express more than one category) can only be inserted where multiple terminal nodes have been joined or fused by independent principles. I show that no such independent principles can be motivated for French fused preposition-determiners, which, I argue, are fused only because portmanteau morphemes for them exist. This possibility follows naturally from phase-sized lexical insertion.

Enlarging the domain of lexical insertion operations, without any other changes, would have the effect of relaxing the restrictiveness of the model, compared with standard DM. However, I recapture a high degree of restrictiveness by showing that lexical insertion proceeds in two stages, a syntactic stage (called L-MATCH in Bye and Svenonius 2010; in press) and a phonological stage (called INSERT, op cit.), with strict modular encapsulation of the two stages (in accordance with the PPFS and the IRS). This means that 'normal' phonology takes over an increased role in allomorph selection (compared to standard DM, and in particular the model proposed in Embick 2010), reducing the role of special morphophonological rules or constraints.

2   The size of the cyclic domain for lexical insertion

In those strains of DM which adopt phases, such as Embick (2010), there is a tension between the earlier, morpheme-by-morpheme cycle (Siegel 1974, Allen 1978) and the larger, phase-based cycle. I argue that cyclic effects are largely due to the phase-cycle (hence more like Kiparsky's 1982 levels). One result is the possibility of outward sensitivity to phonology.

2.1 Outward sensitivity to phonology

The different kinds of 'inward' and 'outward' sensitivity, from the perspective of the root, have been discussed most extensively by Carstairs-McCarthy (Carstairs 1987; 1988; 1990, Carstairs-McCarthy 2001; 2003). Here I address the directionality of sensitivity to phonological features, not morphosyntactic ones. It is uncontroversial that allomorphy can be 'inward' sensitive to phonology, as in the well-known cases of the Korean nominative case in (1a), with /i/ after consonants and /ka/ after vowels, or the Haitian Creole definite suffix in (1b), where /la/ appears after consonants and /a/ after vowels (focusing on examples which clearly involve suppletion; examples in which the two allomorphs could be derived from a common source raise distinct questions).

(1)   a.   Korean: *pap-i* 'cooked rice-NOM'; *ai-ka* 'child-NOM'
          b.   Haitian Creole: *liv-la* 'the book'; *tu-a* 'the hole'

There are two controversies. One is what role phonology plays in cases like those in (1). One possibility is that the phonological component selects the allomorph, e.g. by syllable well-formedness in (1a). Embick (2010) calls this possibility PHONOLOGICAL SELECTION (PS). He points out that it is problematic for cases like (1b), and would require some morphological information to be somehow visible in the phonological component (as in Bonet et al's 2007 solution in terms of PRIORITY). The other possibility is that the lexical entry for one or both allomorphs includes a kind of contextual condition, which makes reference to phonological properties of the host, but is not itself part of the phonological component. Embick calls this PHONOLOGICALLY CONDITIONED ALLOMORPHY (PCA), and argues that it is the right solution not only for (1b), but also for (1a). I accept Embick's arguments that PCA exists for cases like (1b), but argue below that PS exists as well (and thus might be right for situations like (1a)).

The second controversy is whether phonological sensitivity is ever outward. Bobaljik (2000) has argued that it is never outward, and that this supports the idea that exponents are inserted into the syntactic tree node by node from the bottom upward. I argue here that there is outward sensitivity, which follows if the cycle for lexical insertion is the phase, not the terminal node.

Carstairs-McCarthy (Carstairs 1987, 1988, 1990) provides several examples of putative outward-sensitivity to phonology, but each one is potentially subject to an alternative analysis, in a DM framework, either because they involve alternations in the root, or because the range of environments is highly restricted.

For example, Carstairs-McCarthy's main example of root allomorphy is Italian irregular preterit stems: *romp-* 'break,' *mov-* 'move,' *prend-* 'take,' etc. when unstressed; *rupp-*, *moss-*, *pres-* etc. when stressed. Since stress is only determined after suffixes are in place, this counts as outward-sensitive. However, the theory of DM as proposed by Halle and Marantz (1993) rejects the possibility of root suppletion, using powerful readjustment rules to change the form of roots. So for that theory, these examples could not be PCA.

Carstairs-McCarthy's (1987:187f) clearest example involving an affix is the Fulfulde General Future Active tense, with a suffix /Vt/ (V an underspecified vowel) before V-initial suffixes and /ay/ before consonant-initial ones. However, judging from his description, there are only six possible suffixes (person×number) which can follow the tense morpheme; the first and second person singular are vowel-initial and the rest are consonant-initial. It cannot be positively ruled out that the /Vt/ allomorph is selected in the presence of first and second person singular.

So in the end, we have very few clear cases of outward-sensitive phonological allomorphy (whether PS or PCA). I present a new one here, which clearly involves listing (because there is no plausible derivation between /m/ and /ð/ in Icelandic) and clearly involves phonological conditioning (because the conditioning paradigm is relatively large, and the shape of the conditioning context clearly follows phonological and not morphosyntactic featural lines).

2.2 Icelandic mixed participles

Icelandic, like other Germanic languages, has a complex set of conjugation classes with a basic division into strong and weak verbs. Strong verbs take different endings from weak verbs, have characteristic vowel changes in the root, and form passive and perfect participles with a suffix /m/. Weak verbs take a participial ending /ð/ (compare English *broken* versus *smashed*).

The two classes are illustrated here. Icelandic participles inflect for gender, number, and case, for a paradigm of twenty-four cells (reduced by systematic syncretisms to sixteen independent forms).

(2) The paradigm for weak verb participles, exemplified by *smyrja* 'smear, butter, lubricate'

	M SG	**F** SG	**N** SG	**M** PL	**F** PL	**N** PL
NOM	smyr-ð-ʏr	smyr-ð	smyr-t	smyr-ð-ɪr	smyr-ð-ar	smyr-ð
ACC	smyr-ð-an	smyr-ð-a	smyr-t	smyr-ð-a	smyr-ð-ar	smyr-ð
DAT	smyr-ð-ʏm	smyr-ð-ri	smyr-ð-ʏ	smyr-ð-ʏm	smyr-ð-ʏm	smyr-ð-ʏm
GEN	smyr-ð-s	smyr-ð-rar	smyr-ð-s	smyr-ð-ra	smyr-ð-ra	smyr-ð-ra

(3) The paradigm for strong verb participles, exemplified by *flýja* 'flee'

	M SG	**F** SG	**N** SG	**M** PL	**F** PL	**N** PL
NOM	flu:-m-n	flu:-m	flu:-ɪð	flu:-n-ɪr	flu:-n-ar	flu:-m
ACC	flu:-mn	flu:-n-a	flu:-ɪð	flu:-n-a	flu:-n-ar	flu:-m
DAT	flu:-n-ʏm	flu:-m-nɪ	flu:-n-ʏ	flu:-n-ʏm	flu:-n-ʏm	flu:-n-ʏm
GEN	flu:-m-s	flu:-m-nar	flu:-m-s	flu:-m-na	flu:-m-na	flu:-m-na

The agreement endings following the participial suffix are regularly drawn from the adjectival declension and are fully productive. The overall pattern is clear: the weak verbs have /ð/ where the strong verbs have /m/, and are followed by the same set of agreement suffixes. Some surface differences are caused by regular Icelandic phonology: assimilation of /r/ to a preceding /n/, assimilation of /ð/ to a following /t/, deletion of unstressed vowels (here, /ɪ/) in open syllables before weak endings, and epenthesis of /ʏ/ to break up an illicit word-final /ðr/ cluster (Kiparsky 1984, Rögnvaldsson 1986, Árnason 2005). There are two portmanteaux in the strong paradigm, in the accusative singular of the masculine and in the nominative/accusative singular of the neuter. Svenonius (2012a) works out the morphophonological alternations in detail.

In addition to over a hundred strong verbs and an open class of weak verbs adhering to the above generalizations, there is a class of several dozen verbs which are generally weak in their conjugation, but which show a vowel change in their stem (cf. English *sleep-slept*). In Icelandic, these verbs show a 'mixed' pattern in participle formation, employing both /ð/ and /m/:

(4) The paradigm for mixed participles, exemplified by *berja* 'beat'

	M SG	**F** SG	**N** SG	**M** PL	**F** PL	**N** PL
NOM	ba:r-m-n	ba:r-m	ba:r-ɪð	bar-ð-ɪr	bar-ð-ar	ba:r-m
ACC	ba:r-mn	bar-ð-a	ba:r-ɪð	bar-ð-a	bar-ð-ar	ba:r-m
DAT	bør-ð-ʏm	ba:r-m-nɪ	bør-ð-u	bør-ð-ʏm	bør-ð-ʏm	bør-ð-ʏm
GEN	ba:r-m-s	ba:r-m-nar	ba:r-m-s	ba:r-m-na	ba:r-m-na	ba:r-m-na

The distribution of the participial suffixes is clearly phonological: /ð/ is used before an agreement suffix which is vowel-initial, /m/ elsewhere (and the strong portmanteaux are employed where available). I argue that independently motivated considerations of Icelandic phonology lead to this distribution, once conjugation class is left out of the picture, and argue that it is a case of PS, not PCA. I also show how independent considerations explain why the strong portmanteaux appear in the mixed declension.

This example clearly manifests outward-sensitive phonological determination of allomorphy. Unlike Carstairs-McCarthy's Italian examples, it does not involve changes in the root, so cannot

be treated by any appeal to the special status of roots. And in contrast to Carstairs-McCarthy's Fulfulde example, the paradigm is large enough that it can be seen that there is no plausible set of morphosyntactic features which could be responsible for the alternation.

3 The nature of the target of lexical insertion

Further support for the model is provided by French P–D fusion, where lexical insertion does not target terminal nodes, but sequences of adjacent terminal nodes. This could reflect a syntactic phenomenon, were it not for the interaction with phonology: the fused P-D is blocked in case the noun phrase is vowel-initial.

In French, the definite article has three surface forms in the singular.

(5)   a. *la fille*   'the girl (F)'   c. *l'échelle*   'the ladder (F)'
       b. *le garçon*   'the boy (M)'   d. *l'échelon*   'the rung (M)'

These three surface forms can easily be derived from two underlying forms by regular principles of French phonology (Charette 1991, van Oostendorp 2000), since schwa is regularly deleted before a vowel. If the form /lə/ is underspecified for gender, then phonological considerations will prefer it before a vowel, even for feminine nouns.

French has some fused preposition-determiners, for example instead of the expected *à le* /alə/, we find *au* /o/, and instead of *de le* /dələ/ we find *du* /dy/, as seen in the table below (from Zwicky 1987:212).

(6)

	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*
	/dələkɔl/	/dəlamɛzɔ̃/	/dələpital/	/dypark/

Despite the existence of fused forms, there is every reason to think that P and D are distinct syntactic heads in French. The syntactic distribution of *la femme* is essentially as expected if it lacks a P node (occurring, for example, in argument positions and not generally in adjunct positions), while the distribution of *à la femme* is essentially as expected if it has a P node (occurring in modifier positions). The two can be separated, and the fusion is blocked if a modifier occurs between P and D, as in *de tout le parc*, as noted by Mascaró (1996).

In a model with late insertion starting from the bottom up, the fusion means that the exponent of the French determiner cannot be inserted before the higher context is inspected, to see whether it is one of the fusable P's. Such a state of affairs is not unusual in syntax. For example, a correct case form of a noun cannot be selected in case languages until the context surrounding the DP is inspected. This can be handled by feature transmission in the syntax, before lexical insertion. In principle, a language could have a specially marked P in the syntax which causes a lower D to incorporate. However, what makes the French example special in this respect is the sensitivity to phonology. The phonology of the following noun phrase interferes with the incorporation, as seen in the masculine V-initial forms in (6).

This is bad news for the incorporating P story. There is no good way, in a framework with a strict separation between syntax and phonology, to allow P to syntactically incorporate D just in case D is not phonologically dependent on the following N.

Embick (2007; 2010) proposes a concatenation rule for D which makes reference to phonology, and then another concatenation rule for P which makes reference to nonphonological

information, such as gender and number and whatever features distinguish *de* and *à* from other prepositions. The first concatenation rule bleeds the second.

(7)   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

But this mixes phonology and syntax in a way that would allow all sorts of unattested kinds of interactions. Zwicky and Pullum (1986) and Pullum and Zwicky (1988) argue at some length that syntax makes no reference to phonology, and there is good reason to think that phonology in turn makes only extremely limited reference to syntax, formalized for example in Inkelas' (1989) Indirect Reference principle, which states that phonological rules can only refer to syntactic structure via prosodic constituency. Embick's rules of concatenation make apparently unprincipled reference to both syntax and phonology in addition to positing the ad hoc diacritic '+,' which picks out *de* and *à* on P and picks out masculine and plural on D.

In the framework that I propose here, the lexical entries entirely drive the distribution of fused forms, with no need for an independent operation of fusion. A portmanteau lexical entry can be inserted under two heads at once as long as the two heads are included in the same phase, something which is motivated by the widespread incidence crosslinguistically of P–D interactions.

Specifically, L-Match sees both P and D in the same cycle, and matches both fused and nonfused alternatives for masculine nouns (feminine nouns will not be L-Matched to the fused alternatives because those are specified as masculine). Portmanteau lexical entries are needed on any account, and on this account it follows from their existence that they may be inserted, since insertion is not restricted to terminal nodes (the central tenet of Nanosyntax, Starke 2009).

In the phonological component, the fused and nonfused alternatives compete. Fused forms have a systematic advantage over nonfused forms in that they introduce less prosodic structure. An independently motivated constraint preferring less to more structure (*STRUCT) then gives an edge to the portmanteau forms. However, hiatus avoidance (ONSET) outranks the avoidance of structure, in French, and so if the noun phrase is vowel-initial, the version with the defective vowel in the determiner is preferred. A draft of this analysis is presented in Svenonius (2012b).

4   Conclusion

I have argued for an explicit theory of lexical insertion which combines important features of DM and Nanosyntax. It relies on a simplified version of the cycle, compared to DM, by expanding the domain of lexical insertion, following Nanosyntax, and applying this to the phase. It achieves a strict separation of syntax and phonology by introducing a two-stage process of lexical insertion.

Linguistic theory of the twentieth century was based on a small number of syntactic categories, each of which bore a large number of features. These feature-laden heads were the locus of many operations. In the twenty-first century, features have been increasingly parceled out into separate heads, increasing the height of the trees but simplifying their leaves. This has allowed for a more explicit and well-grounded theory of featural interactions, but has lost an account of some of the phenomena that were previously associated with the domain of the syntactic category. The phase reintroduces a unit which groups together a number of distinct features, and so in a way is like head of the twentieth-century. So what was a head is now, roughly, a phase, and what was a feature is now, roughly, a head. This is not exactly right, since e.g. P and D in French were not features on a single head in the twentieth century, but there seems to be some truth to it.

References

Allen, Margaret. 1978. *Morphological Investigations*. Ph.D. thesis, University of Connecticut, Storrs, CT.

Árnason, Kristján. 2005. *Hljóð: Handbók um Hljóðfræði og Hljóðkerfisfræði*, vol. I of *Íslensk Tunga*. Almenna Bókfélagið, Reykjavík.

Bobaljik, Jonathan David. 2000. The ins and outs of contextual allomorphy. In *University of Maryland Working Papers in Linguistics*, edited by K.K. Grohmann and C. Struijke, vol. 10, pp. 35–71. UMCP.

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

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. OUP, Oxford.

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

Carstairs, Andrew. 1987. *Allomorphy in Inflexion*. Croom Helm, London.

Carstairs, Andrew. 1988. Some implications of phonologically conditioned suppletion. *Yearbook of Morphology* 1988: 67–94. Reprinted in 2001 in *Phonology: Critical Concepts*, vol. V, ed. by Charles S. Kreidler, 111-139, Routledge, London.

Carstairs, Andrew. 1990. Phonologically conditioned suppletion. In *Contemporary Morphology*, edited by Wolfgang U. Dressler et al., pp. 17–23. Mouton de Gruyter, Berlin.

Carstairs-McCarthy, Andrew. 2001. Grammatically conditioned allomorphy, paradigmatic structure, and the ancestry constraint. *Transactions of the Philological Society* 99: 223–245.

Carstairs-McCarthy, Andrew. 2003. Directionality and locality in allomorphy: A response to Adger, Béjar and Harbour. *Transactions of the Philological Society* 101 1: 117–124.

Charette, Monik. 1991. *Conditions on Phonological Government*. Cambridge University Press, Cambridge.

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.

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

Embick, David and Rolf Noyer. 2007. Distributed Morphology and the syntax/morphology interface. In *The Handbook of Linguistic Interfaces*, edited by Gillian Ramchand and Charles Reiss. OUP, Oxford.

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

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The View from Building 20*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 111–176. MIT Press, Cambridge, Ma.

Harley, Heidi and Rolf Noyer. 1999. Distributed Morphology. *Glot International* 4 4: 3–9.

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

Kiparsky, Paul. 1973. 'Elsewhere' in phonology. In *A Festschrift for Morris Halle*, edited by Paul Kiparsky and Steven Anderson. Holt, Rinehart and Winston, New York.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In *Linguistics in the Morning Calm*, edited by I. S. Yang, pp. 1–91. The Linguistic Society of Korea and Hanshin, Seoul.

Kiparsky, Paul. 1984. On the lexical phonology of Icelandic. In *Nordic Prosody III: Papers from a Symposium*, edited by Claes-Christian Elert, Iréne Johansson, and Eva Strangert, pp. 135–164. Almqvist & Wiksell International for University of Umeå, Stockholm.

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

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

Pantcheva, Marina. 2011. *Decomposing Path: The Nanosyntax of Directional Expressions*. Ph.D. thesis, University of Tromsø.

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.

Rögnvaldsson, Eiríkur. 1986. *Íslensk orðhlutafræði: Kennslukver handa nemendum á háskólastigi*. Háskóla Íslands, Reykjavík.

Siegel, Dorothy. 1974. *Topics in English Morphology*. Ph.D. thesis, MIT, Cambridge, Ma.

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. In Svenonius et al. 2009: 1–6.

Svenonius, Peter. 2012a. Look both ways: Outward-looking allomorphy in Icelandic participles. Ms. University of Tromsø; available at ling.auf.net/lingBuzz/001519.

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

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

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.