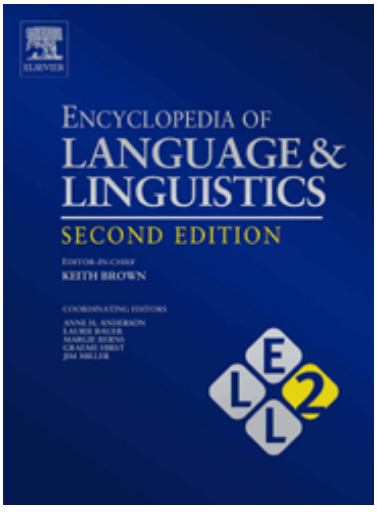

Noyer R (2006), Distributed Morphology. In: Keith Brown, (Editor-in-Chief) *Encyclopedia of Language & Linguistics, Second Edition*, volume 3, pp. 734-737. Oxford: Elsevier.

Distributed Morphology

R Noyer, University of Pennsylvania, Philadelphia, PA, USA

Distributed Morphology, a theory of the architecture of grammar developed in the early 1990s (Halle, 1991; Bonet, 1991; Noyer, 1992; Harris, 1992; Halle and Marantz 1993, 1994), differs from most other approaches to morphology by proposing that words and their subparts are simply syntactic constituents, smaller than but fundamentally no different from larger constituents such as phrases and sentences: although each size of constituent has certain unique properties, all have their internal structures built by the same grammatical component, syntax. Because the lexicon, as it has come to be typically viewed in generative grammar, does not exist within Distributed Morphology, the functions ordinarily attributed to the Lexicon – word-formation, storage of unpredictable sound-meaning correspondences and of morphologically irregular forms, generation of the atomic units of syntactic analysis and so on – are instead *distributed* among various other components of the grammar.

Architecture of Grammar

Within Distributed Morphology the grammar is divided into two parts. First, several distinct repositories contain listed information: a morpheme list, a vocabulary, and an encyclopedia. Second, a generative engine consisting of the syntax proper and various post-syntactic mechanisms (lowering, impoverishment, ornamentation, linearization, and local dislocation) is responsible for building structured linguistic expressions from morphemes chosen from the morpheme list, and interpreting these expressions both phonologically and semantically with information supplied by the vocabulary and the encyclopedia. A schematic view of the model is shown in Figure 1.

The morphemes in the morpheme list form the terminals of morphosyntactic structure; they contain no phonological features, and in this sense are quite different from morphemes as conventionally defined. Morphemes are of two types: Root, which represents an open class item of indeterminate category whose categorial features are determined by its syntactic context (Marantz, 1997), and various others representing functional categories of syntax, such as Tense, v, C, or D, further elaborated by morphosyntactic features (1). It is left to vocabulary items to relate phonological exponents to morphemes and to detail the contextual conditions on the insertion of these exponents (2). Finally, encyclopedia entries relate interpretations and structured linguistic expressions, which may be words or phrases (3). In this way, Distributed Morphology does not recognize the word as a privileged domain for idiosyncratic sound-meaning correspondences; instead phrasal idioms such as *hang ten* 'ride a surfboard' or *bite the dust* 'die' are treated on par with words such as *trans-miss-ion* 'car part.'

(1) Morphemes
  a. [Root]
  b. $\begin{bmatrix}\text{Tense}\\ +\text{past}\end{bmatrix}$

(2) Vocabulary items
  a. /z/ ↔ [+plural]
  b. /d/ ↔ $\begin{bmatrix}\text{Tense}\\ +\text{past}\end{bmatrix}$
  c. /lɛg/ ↔ Root

(3) Encyclopedia entries
  a. [[lɛg $_{\text{Root}}$] $_{\text{n}}$] ↔ 'leg'
  b. [[hæŋ $_{\text{Root}}$] tɛn $_{\text{RootP}}$] ↔ 'ride on a surfboard'

Key Components

Halle and Marantz (1994) identify three key components of the Distributed Morphology model: late insertion, underspecification, and hierarchical structure all the way down.

Late insertion permits phonological exponents to be supplied to a linguistic expression late in its derivation, after syntactic movement has occurred. In other words, syntactic structures are built from and manipulate morphemes whose phonological expression is later supplied by the process of vocabulary insertion (or spell-out). The exact timing of vocabulary insertion is a subject of current debate; but the

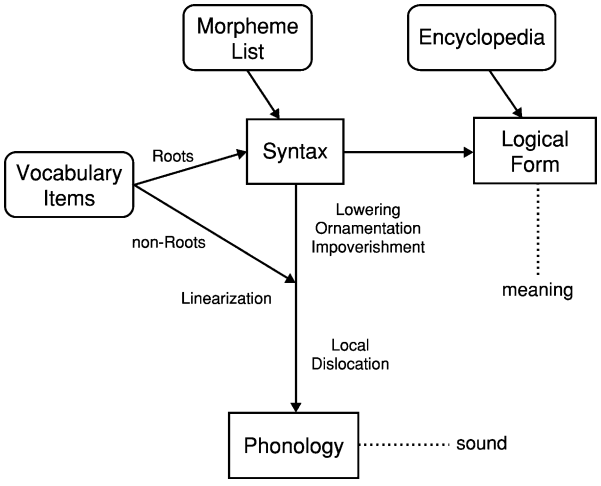

Figure 1 Schematic view of the model.

simplest view, following Embick (2000), is that the exponents of Root morphemes are inserted early, prior to or perhaps cyclically during syntax, while the exponents of other morphemes are supplied after syntax. In this respect Distributed Morphology shares the interpretive approach of other theories, such as the Extended Word and Paradigm model (Anderson, 1982) and Lexeme-Morpheme Base Morphology (Beard, 1995).

Underspecification prevents items of vocabulary from needing to be fully specified relative to their contexts of occurrence, ensuring simple treatments of default exponents. For example, the German adjectival suffix *-n* appears in a quite heterogeneous set of categories most simply described as the contexts where no other affix's more specific conditions for insertion are met. A general principle requiring that the most specific among competing exponents is inserted permits such default exponents to be minimally specified as 'elsewheres' (Lumsden, 1987; Halle, 2000).

An important consequence of late insertion is that features such as number or tense are supplied to the syntactic computation by abstract morphemes that are fully-specified for such features. In lexicalist approaches syntactic features must instead be supplied by words, which in turn obtain these features either by percolation from the specific exponents they contain, or by default rules filling in unmarked values. But because the distribution of default affixes entails that they must be left underspecified, it is predicted that default affixes should never need to supply marked values, i.e., values that cannot be provided by redundancy rules. This prediction is, however, quite problematic, as shown, for example, by the present-tense inflection of English verbs. The affix /-z/ appears only in the third-person singular, all other person and number combinations having a zero-suffix instead. Suppose then that the lexicon contains affixes roughly like the following:

(4) a. $\begin{bmatrix} /z/ \\ \text{present} \\ 3 \\ \text{sg} \end{bmatrix}$

b. $\begin{bmatrix} \emptyset \\ \text{present} \end{bmatrix}$

The lexicon will now generate two present-tense verbs, one marked third-person singular and having the suffix *-z*, the other having no overt suffix and underspecified for person and number. As Noyer (2001) argues, the problem that emerges is, assuming that syntactic structures are projected from freely chosen combinations of lexical items, there is now no obvious way to prevent the verb form unmarked for number from composing with a third-person singular subject. In the derivation of such a sentence, the fact that the lexicon *can* generate a more specified verb form cannot be taken into consideration, since the derivation examines the compatibility of the lexical items that are chosen to project structure, not the range of all possible lexical items that *might have been* chosen to project structure (this set being, in fact, infinite).

To address this problem, lexicalist theories require that words generated by the lexicon do not in fact freely project syntactic structures, but instead must pass through a filtering mechanism that assembles words into paradigm groups, allowing only one word into each paradigm slot (Wunderlich, 1996). Distributed Morphology rejects such filters as redundant, given that the same space of featural possibilities will be generated in the syntax regardless.

Finally, hierarchical structure is assumed to exist at all levels of structural analysis; specifically, there are no 'process morphemes,' and a strict separation is maintained between morphophonological allomorphy rules, which may change the phonological properties of an exponent, and vocabulary insertion, which merely supplies exponents. For example, the derivation of *feet* involves the insertion of the vocabulary item /fʊt/ in a Root position in the context of a plural morpheme, insertion of a zero exponent into the plural morpheme, and finally, a morphophonological readjustment of the stem, changing its syllable nucleus to /ē/. On this view, then, it is incorrect to assert that either /fēt/ or the process of changing /ʊ/ to /ē/ actually spells out [+plural]. Properly speaking, /fēt/ is instead the stem allomorph of /fʊt/ which is generated in the plural environment. The 'piece-based' stance and rejection of processual rules of exponence aligns Distributed Morphology with certain lexicalist theories, most particularly that of Lieber (1981).

Operations after Syntax

Distributed Morphology predicts that in the default instance, the internal structure of words follows from syntactic operations, most particularly head movement (or X^0 Movement). Syntactic word/formation explains the 'mirroring' of syntactic derivational history and morpheme ordering noted by Baker (1985) without appeal to an extra-syntactic mirror principle. Nevertheless, much evidence suggests that in order to derive the full complexity of morpheme ordering the very general movement operations ordinarily imputed to syntax proper must be supplemented by several more specific post-syntactic mechanisms.

Morphological merger (Marantz, 1988) permits a morpheme to exchange a relation of linear adjacency

to another constituent (indicated by * below) for the relation of adjunction to the head of that constituent. In (5a), for example, the Latin second-position clitic *-que* 'and' is initially linearized left-adjacent to the constituent *bonī puerī* 'good boys':

(5) a. [puellae * [que * [bonī * puerī]]]
girls and good boys
b. [puellae * [[bonī+que * puerī]]]
girls good+and boys

In (5b), after Merger has applied, *-que* has traded its relation of left-adjacency to *bonī puerī* for a relation of (right-)adjunction to *bonī*, the peripheral morpheme of this constituent.

Embick and Noyer (2001) distinguish two varieties of merger: lowering and local dislocation. Lowering, such as T-to-v in English (6), is indifferent to linear ordering properties, and adjoins the merged morpheme to the syntactic head of a constituent; it applies prior to vocabulary insertion and is thus insensitive to the idiosyncrasies of specific exponents.

(6) Mary [$_{TP}$ t_i [$_{vP}$ loudly play-TENSE:PAST$_i$ the trumpet]]

Local Dislocation, such as the movement of Latin *-que*, on the other hand, applies after vocabulary insertion and linearization, and adjoins the merged morpheme to the peripheral morpheme of a neighboring constituent.

A second complication arises from affixes that do not plausibly spell out syntactic projections but rather appear to be inserted in specific morphosyntactic contexts. For example, morphological case marking is typically triggered by syntactic configurations and specific case-assigning verbs or prepositions, with no evidence of a syntactically active case morpheme *per se* (Marantz, 1992; McFadden, 2004); Embick (1997) makes similar claims for 'reflexive' verbal clitics such as *se, si* in the Romance languages. These dissociated morphemes are inessential to the operations of syntax, and, inasmuch as their conditions for insertion are already present in syntax, irrelevant to semantic interpretation. Instead, it is hypothesized that they are inserted after syntax by language-specific rules of ornamentation that account in large part for the wide range of morphological complexity observed in natural language.

Finally, impoverishment derives systematic category neutralizations by deleting certain morphosyntactic feature values prior to vocabulary insertion. For example, while classical Arabic lacks a first-person dual category, both in pronouns and in all types of verb agreement, related South Arabian languages Mehri and Soqotri do have this category:

(7) Prefix conjugation agreement

	Sing.	Dual	Plural	
Arabic	ʔ-—-*u*	(= 1st plural)	*n*-—-*u*	1st person
Mehri	ʔ-—-Ø	ʔ-—-*ō*	*n*-—-Ø	
Arabic	*t*-—-*u*	*t*-—-*ā*	*t*-—-*ūna*	2nd-person masculine
Mehri	*t*-—-Ø	*t*-—-*ō*	*t*-—-*əm*	

In Mehri the 1st person dual is spelled with the first-person prefix /ʔ-/ and the dual suffix /-ō/, but the cognate form in Arabic with dual suffix /-ā/ is not well formed. Since this restriction has nothing to do with the exponents themselves, but rather with a general absence of category, a rule of impoverishment is invoked to delete the value [dual] from first-person arguments. As a result, the dual suffix /-ā/ will never be inserted in the first person.

(8) Arabic dual impoverishment

$$[\text{dual}] \rightarrow \emptyset / \left[\begin{array}{c} +1 \\ \_\_\_\_ \end{array} \right]$$

Further Reading

More detailed overviews of Distributed Morphology can be found in Harley and Noyer (2003) and Embick and Noyer (forthcoming).

Bibliography

Anderson S R (1982). 'Where's morphology?' *Linguistic Inquiry 13,* 571–612.

Baker M (1985). 'The Mirror Principle and morphosyntactic explanation.' *Linguistic Inquiry 16,* 373–416.

Beard R (1995). *Lexeme-morpheme base morphology: a general theory of inflection and word formation.* Albany: State University of New York Press.

Bonet E (1991). 'Morphology after syntax: pronominal clitics in Romance.' Ph.D. diss., MIT.

Embick D (1997). 'Voice and the interfaces of syntax.' Ph.D. diss., University of Pennsylvania.

Embick D (2000). 'Features, syntax and categories in the Latin perfect.' *Linguistic Inquiry 31,* 185–230.

Embick D & Noyer R (2001). 'Movement operations after syntax.' *Linguistic Inquiry 32, 555–595.*

Embick D & Noyer R (forthcoming). 'Distributed morphology and the syntax-morphology interface.' In Reiss C & Ramchand G (eds.) *Handbook on interface research in linguistics*. Oxford: Oxford University Press.

Halle M (1991). 'Latvian declension.' In Booij G & van Marle J (eds.) *Yearbook of morphology.* Dordrecht: Kluwer. 33–47.

Halle M (2000). 'Distributed Morphology: impoverishment and fission.' In Lecarme J, Lowenstamm J & Shlonsky U (eds.) *Research in Afroasiatic grammar: papers from the third conference on Afroasiatic languages, Sophia Antipolis, France.* Amsterdam: Benjamins. 125–149.

Halle M & Marantz A (1993). 'Distributed Morphology and the pieces of inflection.' In *The view from building 20*. Cambridge, MA: MIT Press. 111–176.
Halle M & Marantz A (1994). 'Some key features of Distributed Morphology.' In Carnie A & Harley H (eds.) *MIT working papers in linguistics 21*. Cambridge, MA: MIT Working Papers in Linguistics. 275–288.
Harley H & Noyer R (2003). 'Distributed Morphology.' In Cheng C & Rint S (eds.) *The second Glot International state-of-the-article book: The latest in linguistics*. Mouton de Gruyter. 463–496.
Harris J (1992). 'The form classes of Spanish substantives.' In Booij G & van Marle J (eds.) *Yearbook of morphology 1991*. Dordrecht: Kluwer. 65–88.
Lieber R (1981). 'On the organization of the lexicon.' Ph.D. diss., MIT.
Lumsden J (1987). 'Syntactic features: parametric variation in the history of English.' Ph.D. diss., MIT.
Marantz A (1988). 'Clitics, morphological merger, and the mapping to phonological structure.' In Hammond M & Noonan M (eds.) *Theoretical morphology*. San Diego: Academic Press. 253–270.
Marantz A (1992). 'On case and licensing.' In Ao B, Westphal G & Chae H-R (eds.) *Eastern States Conference on Linguistics 91*. Columbus: Department of Linguistics, Ohio State University. 234–253.
Marantz A (1997). 'No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon.' In Dimitriadis A, Siegel L, Surek-Clark C & Williams A (eds.) *Proceedings of the 21st Penn Linguistics Colloquium*. Philadelphia: UPenn Working Papers in Linguistics. 201–225.
McFadden T (2004). 'The position of morphological case in the derivation: a study on the syntax-morphology interface.' Ph.D. diss., University of Pennsylvania.
Noyer R (1992). 'Features, positions and affixes in autonomous morphological structure.' Ph.D. diss., MIT.
Noyer R (2001). 'Clitic sequences in Nunggubuyu and PF convergence.' *Natural Language & Linguistic Theory 19*, 751–826.
Wunderlich D (1996). 'Minimalist morphology: The role of paradigms.' In Booij G & van Marle J (eds.) *Yearbook of morphology 1995*. Dordrecht: Kluwer. 93–114.

Dixon, Robert M. W. (b. 1939)

A Y Aikhenvald, La Trobe University, Bundoora, Australia

R. M. W. (Bob) Dixon was born in Gloucester on January 25, 1939, and was brought up in Nottingham. After graduating with a degree in mathematics from Oxford University, he was in 1961 appointed Research Fellow in Statistical Linguistics at the University of Edinburgh. From reading the works of Boas and Sapir, Dixon realized that the only way he could master the principles of linguistics was to undertake study of a previously undescribed language. This led to 40 years of steady fieldwork, with 25 trips to the Cairns Rain Forest region of North Queensland, three to the Fijian island of Taveuni, and seven to the jungle village of Casa Nova, in Brazilian Amazonia (the trips varied in extent from 10 months to 2 weeks).

In addition to publishing shorter grammars of three Australian languages (Warrgamay, Nyawaygi, and Mbabaram) whose last one or two speakers he worked with, Dixon has published comprehensive grammars of two Australian languages – *The Dyirbal language of North Queensland* (1972) and *A grammar of Yidiny* (1977) – plus *A grammar of Boumaa Fijian* (1988) and *The Jarawara language of southern Amazonia* (2004). He has also published on Dyirbal songs and kinship, in addition to a thesaurus/dictionary plus texts volume on Yidiny (with similar volumes for Dyirbal in preparation).

Bob Dixon is a zealot, believing that anyone who calls himself a linguist should shoulder responsibility for documenting languages. He has supervised the work of more than 30 graduate students, each of whom has completed a high-quality theoretically informed grammar of a previously undescribed language, thus providing an inestimable addition to linguistic knowledge. These have been written in terms of the established and cumulative framework of linguistics treated as a natural science, which has recently come to be called 'basic linguistic theory.' Dixon makes no secret of the poor opinion he holds of people who receive a salary from a linguistics department but simply play around with formal theories or gather data for typological study by lolling in an armchair in a library. As might be expected, such freely expressed opinions do not make him terribly popular with some members of these groups.

A conviction that the basic business of language is the communication of meaning, and that the primary task of linguists is to understand how this is achieved, has characterized his work on 'the semantic basis of grammar.' Dixon's manifold contributions to typological theory have utilized an inductive methodology. The much-cited paper 'Where have all the