

- Fodor J A (2001). 'Language, thought and compositionality.' *Mind and Language* 16(1), 1–15.
- Fodor J A (2003). *Hume variations*. Oxford: Oxford University Press.
- Fodor J A & Lepore E (2002). *The compositionality papers*. Oxford: Clarendon Press.
- Fodor J A & Pylyshyn Z (1988). 'Connectionism and cognitive architecture: a critical analysis.' *Cognition* 28, 3–71.
- Groenendijk J & Stokhoff M (1991). 'Dynamic predicate logic.' *Linguistics and Philosophy* 14, 39–100.
- Hintikka J & Kulas J (1983). *The game of language*. Dordrecht: Reidel.
- Hodges W (1998). 'Compositionality is not the problem.' *Logic and Philosophy* 6, 7–33.
- Horwich P (1998). *Meaning*. Oxford: Clarendon Press.
- Janssen T M V (1997). 'Compositionality.' In van Benthem J & Meulen A T (eds.) *Handbook of logic and language*. Amsterdam: Elsevier. 417–473.
- McLaughlin B (1993). 'The classicism/connectionism battle to win souls.' *Philosophical Studies* 70, 45–72.
- Montague R (1974). *Formal philosophy: selected papers of Richard Montague*. New Haven: Yale University Press.
- Osherson D N & Smith E E (1981). 'On the adequacy of prototype theory as a theory of concepts.' *Cognition* 9, 35–58.
- Pelletier F J (1994). 'The principle of semantic compositionality.' *Topoi* 13, 11–24.
- Rips L J (1995). 'The current status of research on concept combination.' *Mind and Language* 10(1/2), 72–104.
- Smith E E & Medin D L (1981). *Categories and concepts*. Cambridge: Harvard University Press.
- Smolensky P (1987). 'The constituent structure of mental states: a reply to Fodor and Pylyshyn.' *Southern Journal of Philosophy* 26, 137–160.
- Zadrozny W (1994). 'From compositional to systematic semantics.' *Linguistics and Philosophy* 17, 329–342.

Compound

L Bauer, Victoria University of Wellington,
Wellington, New Zealand

© 2006 Elsevier Ltd. All rights reserved.

Definitions

A compound is usually defined (somewhat paradoxically) as a word that is made up of two other words. This basic definition requires a certain amount of modification, some of it for all languages, some of it for specific languages. For example, there may be more than two 'words' involved in the formation of a compound, though there must be at least two. Word is to be understood in this definition as meaning 'lexeme.' The implication of this is that the forms in which the individual subwords appear may be differently defined in different languages: a citation form in one, a stem in another, a specific compounding form in yet a third, a word form in a fourth. Even this leaves room for a certain amount of disagreement about what a compound is in particular languages. Perhaps the rider should be added that the construction created by the two or more lexemes should not be a normal noncompound phrasal structure of the language: *well done* and *in time* are not compounds.

There appear to be two fundamental approaches to the nature of a compound. The first sees a compound as a particular construction type, an entity with a formal definition. The second views a compound as a lexical unit meeting certain criteria. Because the two overlap to a large extent, it may not be clear in which

of these two senses the term is being used. For example, under either approach *blackbird*, *windmill*, and *combination lock* would count as compounds of English, as would *sky-blue*, *onto*, and *freeze-dry*. But the things we find in everyday headlines (such as *PM backs mercy killings bill*, where the relevant unit is *mercy killings bill*) are not included as compounds by those who view compounds as lexical units on the grounds that they arise through the productive use of syntactic rules, but are included as compounds by those who view compounds as a construction type. The argument in favor of the latter view would be that the major distinction between the lexical-item compounds and the headline compounds is frequency of occurrence and that there is no formal distinction to be drawn between the two groups.

Although this leaves us in the unfortunate situation of not necessarily being able to recognize compounds in a given language, there are a number of criteria that are generally accepted as correlating with compound status, at least to a certain degree. These will be considered immediately below.

Orthographic Criteria

Although orthographic criteria cannot be robust, given the number of languages that still are not written or, if written, have had an orthography imposed by non-native speakers of the language, or, if neither of these, may not indicate word breaks in the orthography at all, nevertheless they are taken as a powerful indicator in those languages for which they are

relevant. Compounds are assumed to show their status by being written as single words. This type of criterion is employed linguistically especially by corpus linguists, who may have no other way of isolating compounds. For such people, it is a clearly a less than ideal default. Unfortunately, it is taken as a serious definition by some outside corpus linguistics.

There are a number of problems with such a criterion. First, in a language like English, there is so much variation in the writing of two-word lexical items that even standard dictionaries (supposed arbiters of the prescribed norm) are unable to agree. *Rainforest*, *rain-forest*, and *rain forest* are all easily attestable, for example. It might be argued, however, that such variation merely shows the progress of an item like *rain forest* from syntactic sequence to lexical item. Unfortunately, such drift is not easily correlated with the relative age of dictionaries showing a particular orthography. More serious, though, are items like *a New York–Los Angeles flight*, which, on strict orthographic criteria, give compounds such as *York–Los*. We should also note that orthographic conventions can change, as where recent orthographic changes have made the writing of noun + noun sequences as two words a rather more common event in Danish than it used to be.

Phonological Criteria

There are a number of phonological ways in which status as single words may be indicated. These may be segmental or suprasegmental.

In English, for example, stress is sometimes taken to be criterial, distinguishing most clearly between examples like *blackbird* and *black bird*. Where noun + noun sequences are concerned, stress is less consistent in English, *church-warden*, for example, being reported with various stress patterns. And for examples like *sky-blue*, stress as the sole criterion would suggest that they are compounds when used attributively (*a sky-blue dress*) and not when used predicatively (*her dress was sky-blue*).

In Danish, the *stød* (or glottalization, corresponding to the use of tones in other Scandinavian varieties) is generally lost in the first element of a compound (Basbøll, 1985).

In Japanese, there is a process known as *rendaku*, whereby the initial consonant of the second element of a certain class of compounds becomes voiced, as in the examples in (1).

- | | | | | |
|-----|-----|-----------|-----------------|-----------|
| (1) | iro | ‘color’ | kami | ‘paper’ → |
| | | irogami | ‘colored paper’ | |
| | e | ‘picture’ | tako | ‘kite’ → |
| | | edako | ‘picture kite’ | |

- | | | | |
|---------|--------------------|------|------------|
| ike | ‘arrange’ | hana | ‘flower’ → |
| ikebana | ‘flower arranging’ | | |

In Sengseng, geminate consonants are permitted within a word, but are always broken up by an epenthetic vowel if they occur adjacent at word boundaries.

Morphological Criteria

It is sometimes claimed that since compounds are single lexemes, the only inflection allowed in them must be the inflection that allows that single word’s functioning in the sentence as a whole to be seen. Accordingly, internal words (words acting as modifiers within the compound) are said to disallow inflections.

There are instances in which this appears to be true. In peninsular Scandinavian, for example, where definiteness is marked by a suffix, no marking for definiteness can occur on the modifying element of a compound.

There are other instances in which this rule appears to hold most of the time, but not all of the time. In English, for example, plural is not usually marked on the modifying elements in compounds so that even a noun like *trousers*, which rarely appears in the singular, has no plural marking in the compound *trouser-press*. However, there is an apparently increasing set of items in which internal plural marking is found: *official secrets act*, *suggestions box*, *weapons inspector*. The example of *mercy killings bill* cited earlier is an instance. Although we might want to argue that plural in English is not entirely inflectional, we appear to have sporadic counterexamples here.

Then there are instances in which this rule clearly does not hold. Sometimes these are sporadic, as in Danish *nytår* (new.NEUT-year), where a clearly inflectional neuter agreement occurs contrary to usual practice. In other languages, though, we may find systematic violations of the rule. Consider, for instance, the Finnish examples in (2) (from Sulkala and Karjalainen, 1992) or the Sanskrit examples in (3) (from Whitney, 1889).

- | | |
|-----|-----------------------|
| (2) | auto-n-ikkuna |
| | car·GEN·window |
| | ‘car window’ |
| | maa-lta-pako |
| | country·ABL·flee.NOM |
| | ‘rural depopulation’ |
| (3) | bhayamkarṭr |
| | fear.ACC.causer |
| | ‘causer of fear’ |
| | divikṣīt |
| | sky.LOC.dwelling |
| | ‘dwelling in the sky’ |

Syntactic Criteria

Syntactic criteria are attempts to find ways of indicating that the compound is being treated in the syntax as a single unit and not as a sequence of two distinct units. This usually means that anaphora cannot pick out the modifying element in a compound, but can in a syntactic phrase. Thus, for English, it is noted that we can say *I thought this house had aluminum windows, not wooden ones* (where *ones* refers to windows, not aluminum windows, and thus shows *aluminum windows* to be made up of two elements), and also *The aluminum windows look good but I'm afraid that it may not be strong enough* (where *it*, in referring to the aluminum rather than to the whole compound, again shows that the two words are separate). In parallel instances with lexical compounds such as *combination lock*, the claim would be that *I wanted a combination lock but they only had Yale ones* and *I installed a combination lock and now I can't remember it* (where *it* refers to the combination) are not possible, indicating the unity of the sequence. Because these tests appear to be robust some of the time, they are hard to argue against, but we can make some comments here. First, intuitions are not always as secure as the tests seem to predict; second, speakers do produce the 'impossible' constructions on occasion, as is shown in (4); third, there are constraints on these uses that have not been fully explored; and fourth, because the tests deal with the degree to which the constituent words of a compound can be recognized, they are again tests of lexicalization.

- (4) Here he found that the Greatest Story Ever Told
had stopped for a tea-break. Dorothy
Horncastle was dispensing mugs of the stuff
from a large copper urn . . .
(Reginald Hill, *Bones and Silence* (London:
Collins, 1990: 275))
I want to give myself a headache by banging it on
the floor! (attested)

Semantic Criteria

Semantic criteria sometimes invoked to indicate that something has become a compound are fundamentally indications of lexicalization. A celebrated example concerns the pair *push-chair* and *wheel-chair*. Although it is possible to push a wheel-chair and a push-chair ('baby buggy') has wheels, the two are uniquely identified by their labels that, therefore, imply a lot more than they state. This specialization of meaning is sometimes taken to indicate compound status.

There are two possible counterarguments here. The first is that lexicalization (or idiomatization) is

something that affects not only noun + noun sequences such as *wheel-chair*, but also syntactic sequences such as *a red herring* or *how do you do?* If *how do you do?* is still a sentence, so *wheel-chair* should have the same status before and after lexicalization. The second is that meaning specialization is not something that comes only with frequent use (and gradual movement from syntax to the lexicon) but something that comes with first use. Downing (1977) provides us with the celebrated attested example of *apple-juice seat*. In the abstract this could mean a number of things, including the seat in which you have to sit to win some apple juice, the seat on which apple juice has been spilled, the seat on which a can of apple juice has been placed, the seat on which I usually sit when I drink apple juice, and so on. The attested meaning of 'seat with a glass of apple juice placed before it' is already a specialization from the large number of potential meanings that construction could have had. That being the case, we could argue that any relevant sequence that has actually been used has already had a specialized meaning. All that may remain in question is the degree of specialization.

The Universality of Compounds

Because of the two types of definition of compound referred to earlier, it is not clear whether or not all languages have compounds. Claims that they have can be found in the literature; so can claims in grammars of individual languages that compounds are not found in that language. The two need not be incompatible if they depend on different types of definition. It may be, for example, that compounds viewed as a construction type are universal, but compounds viewed as lexical entities are not. Because the problem has not been recognized in the literature, it is impossible to be sure.

There is also the problem that in individual languages, things may be called compounds that would not normally be so termed in other languages. For example, Glinert (1989) describes as compounds of Hebrew, things that sound like blends (q.v.) to the reader more used to languages such as English. Discussions of French sometimes refer to units such as *chemin de fer* ('way of iron' = 'railway') or *pomme de terre* ('apple of earth' = 'potato') as compounds. It is clear that these items are lexical items (listemes in the terminology of Di Sciullo and Williams, 1987). The use of the label compound may thus be intended to indicate the lexicalized nature of such constructions (or at least, of some such constructions) and their near-syntactic productivity rather than be intended to attribute any particular structure to them. The use of the label compound in such

instances certainly adds to the confusion surrounding the whole issue.

Perhaps just as confusing is the way in which noun incorporation is sometimes not included under compounding. Incorporation is an important enough and theoretically interesting enough method of word formation to demand separate discussion, as is done in this encyclopedia (see **Incorporation**). But since both compounding and incorporation involve the close binding together of two stems into a new morphosyntactic element, there is also a sense in which the two need to be viewed as related processes. Neoclassical compounds, also dealt with separately here (see **Neoclassical Compounding**), have many of the features of compounds, but are compounded according to borrowed patterns rather than native ones.

The Semantics of Compounds

Speakers of European languages, at least, seem to view compounds made up of two nouns as the prototypical type of compound (although there are languages that appear to prefer verb–verb compounds). A noun–noun compound such as *rain–cloud* is an ideal construction for providing a subcategorization. The element *cloud* (the head element; see below) tells us what kind of entity we are dealing with, and the modifying element (here, *rain*) tells us something about the subtype the compound denotes. Similarly, an adjective–noun compound such as *blackbird* provides not so much a description of the bird (female blackbirds are brown) as a label for a subtype of bird. What is perhaps strange in English is that this subcategorizational use of adjectives need not be restricted to compounds: neither a red squirrel nor red wine is prototypically red, yet their labels contrast with *gray squirrel* and *white wine*, respectively, and thus show a subcategorization of precisely the same type that is found in compounds.

Much is made in the literature of the superficial ambiguity of noun–noun compounds. Although a *hayfever pill* may be intended to relieve hayfever, a *sleeping pill* is intended to induce sleep, a *sugar pill* neither relieves nor induces sugar but contains it, and a *morning-after pill* provides a fourth logical link between the elements of the compound. Several approaches to the descriptive problem posed by such apparent ambiguity have been taken in the literature. In no particular order, these include the following:

- i. relating the various logical links to the meanings of prepositions available in the language or inflectional cases available in the language (so ‘pill against hayfever,’ ‘pill for sleeping,’ and so on)

and assuming that the compound arises through the deletion of such marking;

- ii. relating the various logical links to the syntactic role the elements might play in sentences glossing the link (where ‘the pill relieves hayfever’ and ‘the pill induces sleeping’ would both be of the same subject–predicate type, opposed to ‘the pill is taken on the morning after,’ which would be a subject–adverbial type);
- iii. relating the various logical links to specific predicates that are assumed to be deleted in the course of the syntactic derivation of the compound structure (for example, the actual lexemes *RELIEVE* and *INDUCE* might be considered to be present at some underlying level of analysis, but not at the surface);
- iv. relating the various logical links to a limited set of semantically basic predicates that are deleted in the process of derivation (this solution is similar to the last, but assumes some set of universal Aristotelian categories rather than language-specific lexemes);
- v. some mixture of the above.

Quite apart from the theoretical problems that beset all of these approaches, in the final analysis all of these suggestions fall foul of the fact that there are some compounds that are remarkably resistant to any of these classifications. For example, *spaghetti western* requires some lengthy paraphrase (‘western made in a country that can be characterized by the amount of spaghetti that is consumed there’), and yet is not unique: *goulash communism* seems to reflect precisely the same relationship and the relationship underlying *milk tooth* remains obscure.

A preferable solution may be to see the relationship between the elements not as an ambiguity but as a vagueness and to deny that the specific links between the elements of compounds is strictly grammatical at all. Rather, the specificity that speakers read into the meanings of these compounds can be seen as the result of the lexicalization process (q.v.), starting with the context of first use and becoming more specific with further use. Even this solution has problems associated with it, however. Though the ambiguity or vagueness that has been discussed here is found with one type of compound, there are other types that do not show the same variable meaning relationship. Some of these other types will be discussed below.

Compounds and Headedness

In the last section, the point was made that compounds provide a suitable structure for reflecting

subcategorization. Compounds like *rain-cloud* show a modifier-head structure, with the head denoting the superordinate of the thing denoted by the compound, and the modifying element denoting the important feature for subclassification. Compounds of this type denote hyponyms of their head elements.

Whereas the headedness can be defined in semantic terms like this, it is typically the case that headedness can be used to predict rather more about the structure of the compound. First, we can note that compounds of this type all seem to have a binary structure. Even very complex compounds of the headline type can be broken down into a number of binary compounds embedded within each other. At each division we can distinguish a modifier and a head. Second, we note that in languages that have grammatical gender or different inflectional classes, the head of a compound of this type determines the inflectional class for the compound as a whole. Given a German (Standard German) compound like *Zeitgeist* in which the modifying element (*Zeit* 'time') is grammatically feminine and the head element (*Geist* 'spirit') is grammatically masculine, we can tell that the compound will be grammatically masculine and will make its plural in the same way that *Geist* makes its plural. Though there are some apparent counterexamples to this generalization (*highlight* has the past tense form *highlighted* and never **highlit*, despite the inflectional class of *light*), these occur in very special circumstances (here, for example, the verb *highlight* is derived from the noun *highlight*; it is not a compound verb created by the joining of *high* and *light*; compare also *grandstanded*). Also, the head element tends to carry the inflections for the word as a whole, not the modifying element.

There has been some speculation in the literature as to the regularity of the order of modifier and head in compounds of individual languages. There does not appear to be necessary consistency, with languages like Vietnamese and French showing both orders (consider French *homme-grenouille* 'man frog' = 'frogman' as opposed to *chauvesouris* 'bald mouse' = 'bat'). Even English, which is largely right-headed, does not appear to be exclusively so, as is shown by isolated examples such as *whomever* (where the inflections for the word as a whole are carried on the leftmost element) and *Model T* (which denotes a type of model, not a type of T).

This lack of necessary consistency can make it difficult to determine what is treated as the head element in compounds such as *Alsace-Lorraine*, where the meaning is the addition of the two elements rather than a hyponym of one of the elements.

Compounds like *redwood*, which denote not a type of wood but a tree that has red wood, are sometimes

said not to have a head. Although it is true that *redwood* is not a hyponym of *wood*, but of an unexpressed element *tree*, it is nonetheless the case that *red* is the modifier of *wood* and that *wood* is the element that carries the inflections for the word as a whole. Thus, *redwood* may be seen as a headed structure, just as much as *rain-cloud*. Similar arguments seem to hold with examples such as *breath-taking*, *chimney-sweep*, *trouble-free* where there may be problems in applying the hyponymy criterion in a straightforward way.

Classifying Compounds

Compounds have been classified in a number of ways, none of which appears to be totally satisfactory.

The oldest classification is that provided by the Sanskrit grammarians. This provides four fundamental classes of compound, a classification that continues to be used, in whole or in part, today.

Tatpuruṣa compounds are the type in which there is a clear modifier-head structure, as discussed above. These are sometimes termed 'determinative' compounds. In *rain-cloud*, *rain* determines what kind of cloud is denoted. The Sanskrit grammarians give kharmadharaya compounds as a subtype of tatpuruṣa compounds. In more modern times, the kharmadharaya compounds have been divided into two distinct groups. In the first, we have adjective-noun compounds like *blackbird*. The second type seems very different: it is the compound made up of two elements, each of which refers independently to some facet of the thing denoted by the compound as a whole: *man-servant*, *secretary-treasurer*. This latter type is sometimes confused with the next category, dvandva compounds. Dvandvas denote an entity that is the sum of the entities in the two elements: *Alsace-Lorraine* is an example of this type; in many other languages we find dvandvas meaning 'parents' made up of the words for mother and father (for example, in Kashmiri, Marathi, and Tamil). The third main Sanskrit type of compound is the bahuvrīhi compound. In Sanskrit, these were adjectival in nature, with *bahuvrīhi* being an example of the type and meaning 'having much rice'. In more modern descriptions, this label has been appropriated for compounds like *redwood* discussed above, and they are sometimes called possessive compounds or, in German, *Dickkopfkomposita*, again illustrating the type. The final type in the Sanskrit classification is the avyayibhava compound, which is the label given to adjectival compounds used adverbially. The label has tended to be ignored by modern scholars.

Many more modern classifications of compounds are in effect reinterpretations of the Sanskrit labels.

Table 1 Types of compound whose elements can be understood to be coordinated

Examples	Semantic description	Label
<i>man-servant</i> , <i>writer-director</i> , <i>bitter-sweet</i>	The elements denote different aspects of the same individual	Appositional compounds
<i>Alsace-Lorraine</i> , <i>mother-father</i> (e.g., in Tamil)	The compound denotes a unit made up of the individuals denoted by the elements	Dvandva compounds
Vietnamese <i>bàn ghế</i> 'table chair = furniture'	The elements denote different individuals that act as prototypical members of the set denoted by the compound	Usually seen as dvandvas; may be termed co-hyponymic compounds
<i>London-Los Angeles</i> (<i>flight</i>), <i>Greek-English</i> (<i>dictionary</i>)	The elements denote the extremes of some real or metaphorical journey between two points	May be termed translative compounds
<i>American-Australian</i> (<i>talks</i>)	The elements denote the participants; there is no movement between extremes	May be termed participative compounds

For example, dvandvas are sometimes relabeled as copulative compounds, and the *man-servant* type of kharmadharaya is called an appositional compound. This leads to a proliferation of labels without any particular insights. For example, the confusion between appositional and copulative compounds in the literature appears to arise because both can be glossed by inserting the word 'and' between the elements: a man-servant is a man and a servant; Alsace-Lorraine is made up of Alsace and Lorraine. This masks the original insight by focusing on a different superficial similarity, the apparent coordination. But here we can recognize not only the original distinction and instances that may be hard to classify on the borderline, we can recognize further subtypes as well (see Table 1).

Bloomfield (1935) introduces a different fundamental distinction into the classification of compounds. For him, compounds are basically endocentric (they contain their heads) or exocentric (the true head is unexpressed). The tatpuruṣa compounds are endocentric, as are probably the dvandvas. The bahuvrīhi compounds are exocentric. Comment was made above about the headedness of bahuvrīhi compounds like *redwood*. But there are many other compounds that have no overt head element. English examples like *pickpocket* or their equivalents in Romance languages (e.g., French *garde-robe* 'keep dress' = 'wardrobe') also lack overt heads. One unfortunate result of this is that these have sometimes been classified as bahuvrīhi, which, traditionally, they were not. Rather they are a type not readily encompassed by the Sanskrit terminology. Other exocentric types include *sit-in* (a noun made up of a verb and a preposition), *up-keep* (if this is not to be viewed as a nominalization of a phrasal verb through processes or reordering and stress-change), and compounds like *roll-neck* (*sweater*), *red-brick* (*university*), *go-go* (*dancer*), and *high-rise* (*block*), which appear to create nonprototypical

adjectives although they do not have adjectival heads, and things like *army-fashion*, which are most often used adverbially.

A more recent classification sees all the compounds that have been discussed so far as primary compounds (sometimes misleadingly termed root compounds), in contrast to compounds such as *bus-driver*, which are synthetic compounds (sometimes called verbal compounds or verbal-nexus compounds). Synthetic compounds are built around a verb (in *bus-driver*, the verb is *drive*) with arguments of the verb taking up other structural positions in the verb. So *bus* is the direct object of the verb, and the final *-er* suffix represents the subject or external argument of the verb. There are restrictions on what arguments of the verb can occur in the various positions in a synthetic compound structure. Generally adverbial elements and subject arguments do not occur in the first element of such compounds (**He is a fast-driver of buses*, **Driver-cleaning of buses*), except where the second element is a past participle (*home made*, *self-driven*). Apparent exceptions to this rule are interesting, in that they may help delimit the notion of synthetic compound. For some authorities, synthetic compounds are found only when very productive affixes are involved (*-er*, *-ing*, *-ed*, where English is concerned) so that *speech-synthesizer* is a synthetic compound but *speech-synthesis* is not. For such authorities, a compound such as *insect flight* is not a counterexample to the above generalization, though *consumer spending* is. Some authorities are also unwilling to allow spatial or temporal locative elements as parts of synthetic compounds, but others present wider definitions. An example like *Sunday driver* is this a synthetic compound for some but not for others. Although it is clear that this whole area still needs more work, it is of particular interest since it appears to show the interaction of syntactic and morphological principles in the creation of new lexical items.

The Limits of Compounding

Finally, it must be mentioned that there are some things that look like compounds but are not generally accepted as compounds in the literature.

The most obvious type is the type where the compound is the base of some subsequent process. For example, things that look like compound verbs in English are usually not created by straight compounding but by back-formation (*baby-sit* from *baby-sitter*) or conversion/zero derivation (*to grandstand* from *a grandstand*). There is some equivocation here between the final form and the route by which that final form has been achieved. We have already seen that words like *walk-out* may be viewed as nominalizations of phrasal verbs rather than as compounds.

There are some multilexeme lexical items that appear to derive from the lexicalization of a syntactic structure. English examples include *forget-me-not* and *toad-in-the-hole*. Although these are sometimes called compounds, it seems that their formation is completely distinct from that of the constructions that have been discussed here and that they really belong under a different heading.

The same is true of so-called compounds whose primary motivation appears to be phonetic/phonological: things like *namby-pamby* and *shilly-shally*. Those are dealt with separately in this work, since they are not made up of two lexemes.

See also: Back-Formation; Conversion; Hyponymy and Hyperonymy; Incorporation; Lexicalization; Neoclassical Compounding; Panini; Word.

Bibliography

- Adams V (1973). *An introduction to modern English word-formation*. London: Longman.
- Basbøll H (1985). 'Stød in modern Danish.' *Folia Linguistica* 19, 1–50.
- Bauer L (1978). *The grammar of nominal compounding*. Odense: Odense University Press.
- Bauer L (1979). 'On the need for pragmatics in the study of nominal compounding.' *Journal of Pragmatics* 3, 45–50.
- Bauer L (1983). *English word-formation*. Cambridge, UK: Cambridge University Press.
- Bauer L (1983). 'Stress in compounds: A rejoinder.' *English Studies* 64, 47–53.
- Bauer L (1998). 'When is a sequence of noun + noun a compound in English?' *English Language and Linguistics* 2, 65–86.
- Bauer L (2001). 'Compounding.' In Haspelmath M, König E, Oesterreicher W & Raible W (eds.) *Language universals and language typology*. Berlin/New York: de Gruyter. 695–707.
- Bauer L & Renouf A (2001). 'A corpus-based study of compounding in English.' *Journal of English Linguistics* 29, 101–123.
- Benveniste E (1966). 'Formes nouvelles de la composition nominale.' *Bulletin de la Société Linguistique de Paris* 61, 82–95.
- Bloomfield L (1935). *Language*. London: Allen & Unwin.
- Botha R P (1984). *Morphological mechanisms*. Oxford: Pergamon.
- Brekke H E (1970). *Generative Satzsemantik im System der Englischen Nominalkomposition*. Munich: Wilhelm Fink.
- Carr C T (1939). *Nominal compounds in Germanic*. London: Oxford University Press.
- Darmsteter A (1875). *Formation des mots composés en français*. Paris.
- Di Sciullo A-M & Williams E (1987). *On the definition of word*. Cambridge, MA: MIT Press.
- Downing P (1977). 'On the creation and use of English compound nouns.' *Language* 53, 810–842.
- Fabb N (1998). 'Compounding.' In Spencer A & Zwicky A M (eds.) *The handbook of morphology*. Oxford, UK/ Malden, MA: Blackwell. 66–83.
- Farnetani E, Torsello C T *et al.* (1988). 'English compound versus non-compound noun phrases in discourse: An acoustic and perceptual study.' *Language and Speech* 31, 157–180.
- Gleitman L R & Gleitman H (1970). *Phrase and paraphrase*. New York: Norton.
- Glinert L (1989). *The grammar of Modern Hebrew*. Cambridge, UK: Cambridge University Press.
- Hatcher A G (1960). 'An introduction to the analysis of English noun compounds.' *Word* 16, 356–373.
- Jespersen O (1942). *A modern English grammar on historical principles. Part VI. Morphology*. London: George Allen and Unwin; Copenhagen: Munksgaard.
- Kvam A M (1990). 'Three-part noun combinations in English, composition–meaning–stress.' *English Studies* 71, 152–161.
- Ladd D R (1984). 'English compound stress.' In Gibbon D & Richter H (eds.) *Intonation, accent and rhythm*. Berlin/New York: de Gruyter. 253–266.
- Lees R B (1960). *The grammar of English nominalizations*. The Hague: Mouton.
- Lees R B (1970). 'Problems in the grammatical analysis of English nominal compounds.' In Bierwisch M & Heidolph K E (eds.) *Progress in linguistics*. The Hague: Mouton. 174–186.
- Levi J N (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.
- Lewicka H (1963). 'Réflexions théoriques sur la composition des mots en ancien et en moyen français.' *Kwartalnik Neofilologiczny* 10, 131–149.
- Marchand H (1969). *The categories and types of present-day English word-formation* (2nd edn.). Munich: Beck.
- Meys W J (1975). *Compound adjectives in English and the ideal speaker–listener*. Amsterdam: North Holland.

- Olsen S (2000). 'Composition.' In Booij G, Lehmann C & Mugdan J (eds.) *Morphologie/morphology*, vol. 1. Berlin and New York: de Gruyter. 897–916.
- Rohrer C (1977). *Die Wortzusammensetzung im modernen Französisch*. Tübingen: Narr.
- Ryder M E (1994). *Ordered chaos: The interpretation of English noun–noun compounds*. Berkeley: University of California Press.
- Sandra D (1990). 'On the representation and processing of compound words: Automatic access to constituent morphemes does not occur.' *Quarterly Journal of Experimental Psychology* 42A, 529–567.
- Scalise S (1992). 'The morphology of compounding.' *Italian Journal of Linguistics/Rivista di Linguistica* 4/1.
- Sulkala H & Karjalainen M (1992). *Finnish*. London and New York: Routledge.
- Warren B (1978). *Semantic patterns of noun–noun compounds*. Göteborg: Acta Universitatis Gothoburgensis.
- Whitney W D (1889). *Sanskrit grammar* (2nd edn.). Cambridge, MA: Harvard University Press/London: Oxford University Press.

Computational Approaches to Language Acquisition

J Elman, University of California San Diego, La Jolla, CA, USA

© 2006 Elsevier Ltd. All rights reserved.

Why Computational Models?

One of the striking developments in the field of child language acquisition within the past two decades has been the dramatic increase in the use of computational models as a way of understanding the acquisition process. In part, this has been driven by the widespread availability of inexpensive but powerful computers and the development of software that has made modeling more widely accessible. But there is a more interesting and scientifically significant reason for the phenomena. Computational theories of learning themselves have matured significantly since the middle of the 20th century. The renaissance in neural network (or connectionist) approaches, and more recently, the linkages with mathematical approaches such as Bayesian inference, information theory, and statistical learning have provided a much more sophisticated perspective on a number of issues relevant to language acquisition.

These models are necessarily used in conjunction with empirical approaches, but provide an important complement to such approaches. At the very least, computational models can be thought of as enforcing a level of detail and specificity in a theory or account that a verbal description might not possess. Furthermore, even in simple models, there may be interactions among the model's components that are sufficiently complex that only through empirically running a simulation is it possible to know how the model will behave. Computer models also afford the opportunity to explore aspects of a phenomenon that may not be easily tested in the real world (either because the corresponding situation has not yet been

studied, or perhaps may be infeasible to test). By systematically exploring the full parameter space of a theory, one can sometimes gain insight into the deeper principles that underlie a behavior. And of course, a model may be amenable to analytic techniques that are not practical with real children. With children we can never do more than make inferences about the internal mechanisms that drive a behavior. Computer models, on the other hand, can in principle be completely understood. Finally, such models can serve as hypothesis generators. They often suggest novel ways of understanding a phenomenon. Of course, the validity of the hypothesis ultimately depends on empirical testing with real children.

In general, there have emerged two complementary approaches to modeling. In the first, the goal is to determine that a problem can be solved without making specific claims that the solution implemented in the computer model is the same as it would be for the child. These approaches tend to be more mathematical in nature. Work involving Bayesian inference, information theory, and statistical learning are of this sort. The second approach attempts to model the acquisition process a bit more directly. Learning plays a central role in these approaches, and the models' behavior at intermediate stages is as much a focus as the ability to ultimately master the task. Connectionist models are examples of this second sort of model.

Because the field of computational approaches to language acquisition has grown so explosively – and cannot be fully covered in the present short review – what follows will be organized around the major issues that have been addressed (leaving aside a large number of interesting but less central phenomena). For excellent discussion of related computational approaches, see Brent (1996), Munakata and McClelland (2003), and MacWhinney (1999).