Word-Formation

An International Handbook
of the Languages of Europe

Volume 1

Edited by

Peter O. Müller
Ingeborg Ohnheiser
Susan Olsen
Franz Rainer

De Gruyter Mouton

10. Word-formation in cognitive grammar

Abstract

This article summarizes some of the basic assumptions of cognitive grammar with an eye on their relevance to the study of word-formation. Issues addressed include analyzability, entrenchment, and productivity; the different processes of word-formation (affixation, compounding, and blending); the rationale for recognizing words as distinctive units of linguistic structure; and the similarities and differences between word structure and phrase structure.

1. Basic principles of cognitive grammar

Cognitive grammar is a theory of language which Ronald Langacker began working on in the late 1970s. The standard exposition of the theory is the two-volume *Foundations of Cognitive Grammar* (Langacker 1987, 1992); an updated account is Langacker (2008). Taylor (2002) is a text-book introduction. (For the position of cognitive grammar within a more broadly characterized cognitive linguistics movement, see Taylor 2002, Ch. 1.) Each of these publications has extended discussions of words and their structure. Additional resources − both those which explicitly adopt Langacker's theory and those which are broadly compatible with its goals − are cited in this article.

The starting point of cognitive grammar is uncontroversial: a language is a means for relating sound and meaning. More specifically, a language enables speakers to represent their thoughts and intentions by making available to them an inventory of symbolic associations between units of form (phonological structures) and units of meaning (semantic structures). Hearers familiar with the symbolic associations are able to recover, or to intimate, the speaker's semantic intentions.

A notable feature of the theory, in contrast with competing accounts, is its minimalist approach to the form-meaning relation. There are, namely, only three objects of study in cognitive grammar. The first comprises language in its perceptible form, typically as sound, but also as writing, or (in the case of signed languages) as gesture. Using the term rather broadly, this aspect of language is referred to as phonological structure. The second is the symbolized content, referred to as semantic structure. Semantic structure comprises a person's conceptualizations, in a very general understanding of the term; it includes not only referential intent and propositional content, but also affect, evaluations,

and attitudes towards possible hearers. The third object of study are symbolic associations between phonological and semantic structures. The latter are taken, following Saussure (1964 [1916]), to be conventional in character; their properties and indeed their very existence need to be specifically learned by members of a speech community.

Cognitive grammar proposes a direct relation between phonological and semantic structure. In many alternative theories the relation, on the contrary, is mediated by intervening levels of organization. These are autonomous in the sense that they are composed of elements and relations which are unique to these levels and which cannot be reduced to matters of phonology or semantics. One such intervening level (or, in some theories, cluster of intervening levels) is the syntax. Cognitive grammar does not deny the existence of syntactic phenomena. The claim, rather, is that syntactic elements – notions of lexical and phrasal categories (noun, noun phrase, etc.); categories such as complement and adjunct; syntactic relations (clausal subject, phrasal head, etc.), and even the very notions of word and phrase – can themselves be fully described in terms of the minimalist ontology of the theory. An aim of this article is to indicate the validity of this approach to the description of words and their internal structure.

1.1. A structured inventory of units

In cognitive grammar, knowledge of a language is equated with knowledge of an inventory of units, a unit being defined as any phonological, semantic, or symbolic structure that has been established, or entrenched, in the speaker's mind through frequency of previous use. Importantly, the units do not constitute a random, unordered list. Rather, the inventory is structured by virtue of the ways in which the units are related to each other. Three kinds of relation are of special importance. These are the part-whole relation, the schema-instance relation, and the relation of similarity (or, in the limiting case, identity).

a) *The part-whole relation*. This is the relation whereby a linguistic structure may be analysed into its component parts. The part-whole relation applies to phonological, semantic, and symbolic structures alike. Thus, the phonological form [sɪŋ] breaks down into its component sound units, namely, the sounds [s], [ɪ], and [ŋ], while the symbolic unit [singer] can be analysed into its constituent symbolic units [sing] and [-er]. The counterpart to analysis is composition. This is the process whereby speakers recruit already established units in order to create structures which are not already entrenched in their mental grammar.

b) *The schema-instance relation.* A linguistic structure – whether phonological, semantic, or symbolic – may count as an instance of a more abstractly characterized unit. The phonological structure [sɪŋ] may be seen as an instance of a schematic phonological unit [Syllable], which in turn can be analysed into its schematically characterized parts [Onset] and [Rhyme]. While speakers encounter only instances, the sanctioning schemas are immanent in the instances to the extent that the instances may be recognized as matching the specifications of the schema.

These two relations are recursive. [A] may be a part of [B], which in turn is a part of [C]. Likewise, [X] may be an instance of [Y], which in turn is an instance of [Z]. The

two relations interact in numerous ways. For example, the very possibility of analysing a structure into its parts is constrained by the existence of a schema which sanctions the analysis; at the same time, the analysis strengthens the sanctioning schema. The interplay of the two relations is also evident in the process of composition. The possibility of creatively combining already established units rests not only on the availability of the combining units themselves, but also on the entrenchment of a schema which sanctions the combination.

c) *The similarity relation*. Although difficult to define formally, similarity − or more accurately, speakers' subjective perception of similarity − plays an important role in the emergence of schemas. The fact that [A] and [B] come to be regarded as instances of [C] rests on the prior recognition that [A] and [B] are similar in some respect(s). Moreover, the similarity between [A] and [B] may be perceived to be similar to the way in which [D] and [E] are similar. The commonality between the two cases may give rise to a higher order schema (Nesset 2008). For example, the perceived similarity between word pairs such as *urbane/urbanity*, *insane/insanity*, *profane/profanity* may give rise to a schema representing the alternating phonological forms.

1.2. The autonomy of phonological and semantic structures

In view of the role of language as a means for giving overt expression to a speaker's conceptualizations, it is only to be expected that the study of symbolic relations will constitute the heart of any cognitive grammar investigation. However, the theory also allows for the possibility that phonological and semantic structures exhibit a degree of autonomy vis-à-vis symbolic relations, in the sense that components of these structures need not participate in symbolic associations. Thus, with respect to phonological structure, units such as vowels and consonants, syllables and their parts, do not of themselves have symbolic value. *Hickory-dickory-dock*, of the nursery rhyme, is perfectly well-formed phonologically, in that the expression is composed of established units and is sanctioned by well-entrenched phonological schemas; the expression, however, corresponds to nothing at all at the semantic level. Similarly, there are many elements of semantic structure which are not overtly symbolized by any phonological material. This is very obviously the case with noun-noun compounds (Benczes 2007). There is nothing in the phonological structures of the expressions to indicate that while a *cheese shop* sells cheese, a *barber shop* does not sell barbers (rather, it sells the service which barbers provide) (Jackendoff 2010: 446).

1.3. A usage-based grammar

Cognitive grammar is a strictly bottom-up, or usage-based model of linguistic knowledge and its acquisition (Tomasello 2003). Abstractions (as encapsulated in schemas) are abducted on the basis of acquaintance with their instances. Repeated exposure leads to the progressive entrenchment of instances and of the sanctioning schemas. Entrenchment is, to be sure, a matter of degree; its extent also varies from language user to language user.

It is therefore taken as rather evident that different speakers of the "same" language may have acquired rather different mental representations.

The usage-basis of cognitive grammar entails that a speaker's mental grammar may contain a good deal of redundancy. Expressions which are fully consistent with a sanctioning schema and which could, in principle, be created compositionally, may nevertheless have unit status, provided that they have been entrenched through previous linguistic experience. For example, high frequency plurals may well be stored in the mental grammar alongside the singular forms.

1.4. Motivation

Symbolic relations are conventional, in the sense that they have to be specifically learned (on the basis of experience). This wording is noncommittal with respect to the Saussurian notions of arbitrariness and motivation. There is, to be sure, much in a language which is arbitrary, the fact that the phonological form [sɪŋ] is associated with the semantic concept 'sing' being but one example. Equally, a very great deal in a language is motivated. Motivation is often taken to refer to the role of semantic structure, or conceptualization more generally, in the shaping of phonological and symbolic structures (Panther and Radden 2011). While not discounting this aspect, the proposal that linguistic knowledge constitutes a structured inventory of units offers an alternative perspective, where each unit stands at the hub of a network of relations (part-whole, schema-instance, and similarity) to other units which are already entrenched (to a greater or lesser degree) in the mental grammar (Taylor 2004). Even a phonological form such as [sɪŋ] − whose association with the concept 'sing' may indeed be arbitrary − is motivated by the fact that it conforms with a schema for syllable structure and is made up of units which are well entrenched elsewhere in the language. Language-internal motivation undoubtedly facilitates the learning, storage, and accessing of linguistic units, and testifies, again, to the redundancy inherent in the language system.

2. Word structure: analyzability

Knowledge of a language consists of a network of units, structured by the part-whole, schema-instance, and similarity relations. The focus in this section is on the part-whole relation, with specific reference to words. On what basis it is possible to identify a part of a word?

Observe that the question, as posed, refers to words, that is, to symbolic associations of form and meaning. The fact that [s] is part of the phonological pole of the symbolic unit [sing] is not strictly speaking relevant to word structure. In speaking of word parts we are primarily interested in parts which themselves have symbolic status, whereby a part of the phonological structure of a word can be associated with a semantic value which contributes to the word's semantic structure.

In order for a word-part to be recognized, a number of conditions need to be satisfied, at least in the canonical cases (some non-canonical cases are discussed below). First, we need to be able to establish an identity relation (or, at least, a relation of similarity)

between a piece of a word's phonology and (a piece of) the phonology of at least one other symbolic unit. Second, we need to be able to associate the phonological segments with an identical (or similar) semantics. Third, the postulated semantic value of the word-part needs to feature in the semantic structure of the word itself. Grounds for identifying the word-part will be strengthened in proportion to the number of symbolic units exhibiting the postulated word-part; in the limiting case, the correspondence may be perceived to be accidental or fortuitous, or indeed not noticed at all. Below, I discuss a range of examples, selected so as to illustrate some of the issues which arise when considering the internal structure of words.

2.1. *farmer*

The analyzability of this word into its parts [farm] and [-er] is rather evident. It is instructive to reflect on why this intuition should be so clear. Essentially, we accept this analysis because it fits in with so many other facts about the English language; the analysis, in other words, is very strongly motivated by language-internal relations. First, *farmer* is obviously related to the well-entrenched verb form *farm*, both phonologically (the phonological structure of the verb is fully incorporated into the derived form) and semantically (the derived word invokes the same kind of activity as the base verb). Second, there exists a large number of words for which a comparable analysis is indicated: *singer*, *walker*, *driver*, and countless more. These words characterize a person in terms of what they do, the activity in question being supplied by the verbal element while the agent notion can plausibly be attributed to the suffix. The existence of this cohort of examples, with comparable semantics, enables the abstraction of a schema which captures their commonality, viz. the agentive noun schema $[_N V\text{-er}]$. The schema is supported by its compatibility with a very general pattern of word-formation in English, namely suffixation, whereby the suffix determines the kind of entity that a word refers to while the word-specific semantic content is supplied by the stem. In this sense, the agentive noun schema is itself an instance of a higher order schema for derived words.

2.2. *philosopher* and *barber*

In the canonical case, a word can be exhaustively divided up into constituent symbolic units. The "building block" metaphor (Langacker 1992: 186) does not always apply. Given the high degree of entrenchment of the agentive noun schema (this being a function of the large number of examples on which it is based), we are likely to detect the [-er] suffix in words such as *philosopher* and *grocer*, and perhaps even in *barber*. These words correspond only partially with the agentive noun schema. While the words undoubtedly characterize a person in terms of what they do, English lacks the putative base verbs *(to) philosoph* [fɪ'lɒsəf], *(to) groce*, and *(to) barb*. In the case of *philosopher* we can appeal to the phonological similarity (and even orthographic identity) with the underlined portions of *philosophy* and *philosophical*. Not only this, but the phonological and semantic relatedness of *philosophy* and *philosopher* matches the phonological and se-

mantic relatedness of other word pairs, such as *geography* and *geographer*, *biography* and *biographer*. For *barber*, the situation is less clear, not only because there is no verb *(to) barb*, but also because there are no semantically related words containing the phonological sequence [bɑːb], or phonologically similar sequences.

2.3. *Thursday*

The case of *barber* is not unusual. We often encounter words, a part of which can be identified as a symbolic unit, the remainder of the word, however, lacks this status. Take, as a simple case, the names of the days of the week. These all terminate in [-day], a unit whose semantic contribution to the names is evident. Yet no meaning at all can be attributed to [thurs-], except in the trivial sense that [thurs-] contrasts with [wednes-] and the dangling parts of the other weekday names.

2.4. *perform*

Sometimes the case for analysis may be quite compelling, even in the absence of semantic relatedness. Take the case of (mostly) bisyllabic words of usually abstract meaning. Initial segments include *ad-*, *con-*, *in-*, *per-*, *pro-*, *trans-*; final segments include *-ceive*, *-fer*, *-form*, *-late*, *-tain*. To be sure, not all combinations are possible, though a good many are. The semantic contribution of the constituent units is likely to be opaque to English speakers, even to those who have an inkling as to their etymology. One would be hard pressed to explain the meaning of *perform* on the basis of the putative meanings of *per-* and *-form*.

2.5. *meat*

In spite of the evident semantic relation between *meat* and *eat* − meat is something that you eat − we would probably reject outright an analysis of *meat* in terms of prefixation of the verb *eat* with [m-]. This is because the analysis fails to fit in with other facts about the language. Other words terminating in *-eat* have nothing to do with eating: *heat, bleat, pleat*, etc. Neither are there any other examples of nouns being formed by prefixing a verb with *m-*. There is therefore no basis for proposing the word-formation schema [$_N$m-V]. Moreover, such a schema would conflict with general patterns of English word-formation. While English is certainly not lacking in derivational prefixes (*un-*, *over-*, *out-*, etc.), prefixation by a single consonant is rare. One of the few examples is the negative import of initial *n-*, as in *one*/*none*, *either*/*neither*, *or*/*nor*, the now archaic *aye*/*nay*, and dialectal *owt* ('something')/*nowt* ('nothing').

Even in the case of *meat*, however, we may need to tread with caution. In light of the vagaries of English spelling, English speakers sometimes have recourse of mnemonic aids for distinguishing the different spellings of homophones. Readers may be familiar with the school-room rule for distinguishing the spellings <principle> and <principal>:

<principal>, with an <a>, is an *adjective*. Likewise, one can imagine poor spellers appealing to the semantic relation in order to distinguish <meat> from <meet>. These observations serve as a reminder, not only of the speaker-specific nature of the mental grammar, but also of the role of orthography in the mental representation of words.

2.6. *more*

Even though we would probably not want to recognize *m-* as a meaningful part of *meat*, there are numerous cases where a phonological constituent may be felt to be associated with a semantic value (Rhodes and Lawler 1981). Several words commencing in *m-* are associated with a large number or quantity: *many*, *more*, *most*, *multiple*, *magnitude*, *million*, *myriad*, as well as, perhaps, *magnificence*, *magnanimous*, *mighty*. There are, of course, many words in *m-* which do not share this semantic value, and some, even, which convey the notion of smallness: *minuscule*, *minimal*, *microscopic*. Nevertheless, the association may contribute to the feeling that the phonological form of seemingly unanalysable words is not entirely arbitrary. Take the case of *droop*, discussed next.

2.7. *droop*

This word would probably be regarded as internally simple, yet correspondences and resonances with other words can be detected. Consider words which terminate in *-oop*, such as *hoop*, *loop*, *coop*, *swoop*, *stoop*, and *scoop*. These words invoke the notion of curvature. (Imagine your posture if you are 'cooped up' in an enclosed space.) Now consider words commencing in *dr-*. These include *dry*, *drip*, *drop* (of water), *draft* (of beer), *drain*, *dredge*, *drench*, *drought*, *draw* (water from a well), *drool*, and *dribble*. These have to do with water (or its absence). Remarkably, the meaning of *droop* is almost compositional: curvature (of a plant) through (lack of) moisture.

Often, the proposed semantic value appears to be based in facts of articulation or perception, or in some more vaguely characterized synaesthesic association between sound and meaning (Firth 1930). High front vowels − where there is a small aperture between tongue and palate − tend to be associated with ideas of smallness, while consonant clusters with voiceless plosives may be associated with sudden or jerky movement (*split*, *spike*, *strip*, etc.). The above discussed examples, however, lack this synaesthesic component. By no stretch of the imagination can [dr] be considered a "watery", or even as a "dry" sound. Whatever their status, these conventionalized sound-symbolic relations do, however, set up resonances within the language network and no doubt contribute to speakers' intuitions that words somehow match their meanings.

2.8. *motel*

This word is a blend of *motor* and *hotel*. The initial part of the one word is dovetailed with the final part of the other: [məʊ[t]ɛl], the blend having a semantic value akin to

that of a compound (*motor hotel*). Blending is very different from canonical cases of analyzability, since the sequences [məʊt] and [tɛl] do not in themselves have symbolic status (with respect to the blend in question). Nevertheless, the viability of a blend depends on the possibility of hearers being able to recover the input words, a crucial factor in this regard being the cohort of words containing the phonological strings in question (Gries 2006). It would be inconceivable to propose [lɒt] as a blend of *land* and *yacht*, since initial [l] is hardly likely to activate the word *land*.

3. Entrenchment and drift; countering analyzability

Wheeler and Schumsky (1980) report a remarkable finding. When speakers of English were invited to divide words up into their component parts, about half failed to mark any internal division in the word *baker*. We can imagine speakers being reluctant to mark a boundary in *barber*. But *baker*?

A likely explanation lies in the entrenchment of the word (Hay 2001; Taylor 2012: 131). A feature of entrenchment is that a unit does not need to be assembled (compositionally) from its parts on each occasion of its use, nor do language users need to refer to its parts in order to understand it (Langacker 1987: 59). It already exists as a pre-formed unit and is stored and accessed as such. As a consequence, the unit's internal structure tends to be obscured. Monitoring tasks show that speakers are slower to recognize a word (such as *of*) when it occurs as part of a frequently occurring string (such as *sort of*) than when it is part of a less entrenched sequence (such as *example of*) (Sosa and MacFarlane 2002).

One symptom of the entrenchment of (potentially analysable) units is the phenomenon of drift, whereby an expression takes on phonological and semantic properties in addition to (or even at variance with) its compositional value; the expression, as it were, acquires a life of its own and drifts away from its source. Inflected forms of nouns and verbs typically have their own characteristic distributions. Plural *eyes* distributes in the language quite differently from singular *eye*: the singular tends to be used in idiomatic locutions (*keep an eye on*, etc.), whereas reference to the organ of sight dominates for the plural form (Sinclair 2004). Derived forms often acquire semantic nuances additional to their compositional values. Bybee (1985) cites the example of *dirty*, an adjective whose use need not invoke the notion of 'dirt'. Similarly, *baker*, as an entrenched unit, is not restricted to referring to 'one who bakes'; more commonly, the word is used as the name of a kind of retailer, or retail outlet, which sells certain kinds of baked goods, typically breads, pastries, and cakes. If you buy something 'from the baker', you are not necessarily buying from 'the person who baked' the stuff.

Phonological drift − evidenced by such processes as assimilation, elision, palatalization, consonant lenition, and vowel reduction, sometimes operating cumulatively over lengthy time periods − is liable to obscure the internal make-up of words. Although still evident from the more conservative spelling, the relation of cupboards to cups and boards is probably lost on most English speakers, as is the relation between *state* and *station*. Informal enquiries suggest that many speakers fail to perceive the relatedness of *preside* and *president*, and even of *horizon* and *horizontal*.

4. Characterizing the parts

Derived words − the example of *farmer* is typical − exhibit a marked asymmetry with respect to the properties of their parts (Taylor 2002; Tuggy 1992). Some of the parameters are applicable to the analysis of phrases; others are specific to the domain of word structure.

a) *Autonomy* vs. *dependence*. A unit is autonomous to the extent that it can be conceptualized in and of itself, without reference to other units. Although initially proposed with reference to semantic units, in particular to verbs and their complements (Langacker 1992: 286), the notion also applies to phonology. Stop consonants are phonologically dependent in that we cannot conceive of a stop consonant without invoking some notion of an adjacent vowel-like segment.
   Suffixes, whether derivational or inflectional, are overwhelmingly phonologically dependent. The suffix [-er], being unstressed, absolutely requires as a host a stressed syllable. Prefixes, on the other hand, tend to be phonologically autonomous. Forms such as *over-* (as in *overeat*), *out-* (as in *outperform*), and *anti-* (as in *anticommunist*) in fact exist as phonologically autonomous words. The same, of course, goes for parts of a compound; both *air* and *port* − parts of *airport* − are phonologically autonomous.

b) *Schematic* vs. *contentful.* Some conceptual units, such as those symbolized by *do*, *be*, and *it*, have a highly schematic semantics; others, such as those symbolized by basic level terms such as *walk*, *sing*, and *table*, are more contentful. Prefixes and suffixes are notable for their semantic schematicity while the stems to which they attach supply the bulk of the semantic content. Prefixed *out-* merely refers to some schematically characterized process (whose content is supplied by the stem to which it attaches) which exceeds some expected norm.
   Some affixes are phonologically schematic. The (regular) plural morpheme of English is an alveolar fricative. Whether the sound emerges as voiced or voiceless, and, if voiced, whether preceded by an epenthetic schwa, is determined by properties of the host.

c) *Valence.* The specification of a unit often requires reference to the kinds of items with which it can combine. (The phenomenon is studied in syntax under the rubric of subcategorization.) A transitive verb (by definition) is one which requires a direct object complement. Another way of putting this is to say that the characterization of a unit may make reference to the (schematically characterized) construction in which it needs to occur, in the case in point, the transitive clause construction. Noun-forming *-ness* needs to attach to an adjective, while agentive *-er* attaches to verb stems. The case of *-er* is, to be sure, more complex. Like many symbolic units, the suffix is polysemous (Panther and Thornburg 2001; Ungerer 2007). Thus, on one of its meanings, the suffix can characterize a person in terms of their location, as in *villager*, *islander*, *Londoner*, and *back-bencher*.

d) *Selection* (or *choosiness*). Syntacticians distinguish between subcategorization and selectional restrictions. While *drink* subcategorizes for a nominal object, it semantically selects the name of a liquid. The "choosiness" of word-parts goes well beyond matters of semantic plausibility. There is no semantic reason why *apt* and *correct* should prefer the periphrastic comparatives *more apt* and *more correct* over the in-

flected forms *apter* and *correcter*, while *poor* and *simple* exhibit the opposite tendency, preferring the inflected forms *poorer* and *simpler*. One factor is the phonological properties of the adjectives − those terminating in a consonant cluster, more specifically an obstruent cluster, tend to shun the inflection (Hilpert 2008; Mondorf 2003). Amongst suffixes with roughly comparable semantic values we likewise find large differences in their range of uses. While noun-forming *-ness* is fairly promiscuous, in that it can attach to a very large and possibly open-ended set of adjectival stems, noun-forming *-th*, *-ity*, and *-al* are much more choosy, being virtually restricted to occurring with a closed set of hosts. Factors which motivate these differences are an important site of research (Baayen and Renouf 1996; Bauer 2001).

e) *Profile determinant.* This term refers to that constituent of an expression which determines the kind of entity that the expression refers to (Langacker 1987: 235). (The notion thus corresponds to the traditional notion of head.) Suffixes are profile determinants. The fact that *farmer* refers to a person, or that *farmed* refers to a past-time event, is due to the affix. The case of prefixes is more diverse. *Un-*, *n-*, and *anti-* do not impinge of the semantic type of the referent. *Out-*, on the other hand, imposes a verbal conceptualization; *(to) out-Nixon (Nixon)*.

f) *Coercion*. One unit may influence the phonological shape of a neighbouring unit. The shape of the (regular) plural morpheme in English is determined by the phonological properties of (the final segment of) its host. More complex are the effects of an affix on its host. For example, a feature of adjective-forming *-ic* is that stress is placed on the immediately preceding syllable; this, in combination with other phonological schemas, may trigger changes in vowel quality. Compare *photograph* and *photographic*.

The above parameters allow us to characterize some of the basic categories in morphological theory (Taylor 2002; Tuggy 2005; Ungerer 2007; van Huyssteen 2010). Suffixes (in English) are phonologically dependent, are semantically schematic, are profile determinants, they can be quite choosy with respect to the hosts to which they attach, and can coerce the phonological structure of their hosts. Clitics, such as possessive *s* in English, are phonologically dependent, are non-coercive, and can attach to practically any kind of host (think of examples like *The person I met's new car*, where the clitic attaches to a verb). Freedom of combination is also a defining feature of words, along with such features as phonological autonomy and (in most cases) semantic contentfulness. The approach also accounts for − indeed, leads us to expect − that the boundaries of these categories will be fuzzy, and that some items will be able to "migrate" from one category to another. When unstressed, prepositions have clitic-like properties, while prefixes such as *anti-* and *ex-* can function as independent words, as befits their phonological autonomy. Parts of a blend can also achieve symbolic status. Kemmer (2003) discusses the case of forms in *-erati* (*chatterati*, *culturati*, and, more recently, *twitterati*), based, according to her, on the blend *glitterati* (*glitter + literati*).

5. Composition

The counterpart of analyzability is composition, the process of creating new forms from existing resources. One approach − we might call it the Lego model − views composition

in terms of the "fit" between the parts. The model works tolerably well for syntax (at least, for basic phrase structure configurations). *Tea* matches the subcategorization and selectional requirements of *drink*; hence *drink tea* is an acceptable (grammatical) formation. The Lego model does not always work, even for syntax. By all accounts, *lie* 'to speak untruthfully' is an intransitive verb. Yet in *He lied his way through the interview* the verb takes a direct object. The transitive use of *lie* is imposed by the constructional schema in which the verb occurs, whereby the specifications of the schema override the properties of a constituent unit (Jackendoff 2010). The Lego model is even less appropriate for word-formation. For one thing, some word-formation processes, such as blending, do not consist at all in the fitting together of already entrenched parts. But even when parts can be identified, their combination is associated with all manner of restrictions, idiosyncrasies, and overrides. One cannot, willy-nilly, attach *-al* to a verb stem to form a noun designating an instance of the event. *Cooker* does not refer to a person who cooks (the concept is pre-empted by the noun *cook*). A person who practises geology is not a geologer (*geologist* is the term). Nor, for that matter, can one be sure that attaching /s/ to a singular noun will always generate the acceptable plural form.

5.1. Productivity

It is common in the literature to speak of the productivity of an affix. By this is meant, not so much the number of forms which exhibit the affix, but the potential of the affix for the creation of new forms. By this measure, *-ness* turns out to be a highly productive noun-forming affix, *-ity*, *-th*, and *-al* much less so (Baayen and Lieber 1991). It would be more accurate, perhaps, to speak of the productivity of the schema which sanctions the use of the affix. Productivity is a function of entrenchment, which itself is a function of frequency of occurrence. The relation to frequency is not a simple one, however (Bybee 1995, 2001; Dąbrowska 2004). At issue is not the frequency with which the schema *per se* is instantiated, but the number of different instances on which it is instantiated (in comparison with the number of counter-indications). Word-formation processes which are not particularly frequent in the language at large may nevertheless be productive, provided that they are instantiated on a sufficiently wide range of examples. Nominals in *-ee* are not all that frequent, yet the word-formation process appears to be quite productive, as a consequence of the large number of different types. Equally, processes which are frequently attested may not be productive. The pattern of past tense formation exhibited by the high frequency verbs *do*, *say,* and *have* is not able to be extended to other verbs.

5.2. The power of the schema

In view of its evident inadequacies, for both syntax and word-formation, the Lego model is being abandoned by an increasing number of theorists in favour of an approach which focuses on the sanctioning schemas (or constructions) (cf. article 12 on word-formation in construction grammar; Taylor 2012; Tomasello 2003). As an illustration of the power of schemas in word-formation, consider the example of *hamburger* (Taylor 2002: 293). According to the *Oxford English Dictionary*, the word was used in the late 19th century

to refer to a minced meat patty, a so-called 'Hamburger steak', supposedly a speciality of the city of Hamburg. At this time, the word was presumably analysed as [Hamburg] + [-er]. At some point, however, the word was re-analysed as [ham] + [burger]. We can be confident of this because of the existence of forms such as *cheeseburger*, and even the emergence of *burger* as an autonomous unit.

What could have triggered the reanalysis? Hamburgers, after all, are made of beef, not ham. The crucial factor is not the semantics, but the phonology, more precisely, a phonological schema. The stress pattern of *hamburger* ['σ ˌσ σ] matches the stress pattern of compound nouns; compare *dog-whistle*, *clothes cupboard*, *soup-kitchen*, and countless more. When applied to *hamburger*, the schema indicates a comparable analysis, with [-burger] emerging as the profile determinant and [ham-] as a kind of nominal modifier. The way was then open for *burger* to emerge as an independent word and for it to be modified by other words in a compound structure.

6. Conclusion

The minimalist ontology of cognitive grammar, in association with its rigid adherence to a usage-based account of language acquisition, offers the possibility of a rigorous and constrained description of the internal structure of words and of the processes which sanction creative word-formation. As will be apparent from the brief overview presented in this article, the cognitive grammar approach relies heavily on the notions of entrenchment and of schematic representations; the latter arise through familiarity with instances and are able to sanction both the analyzability of encountered expressions and the creative construction of new expressions. Moreover, the issues raised in this article are by no means unique to the analysis of words and the creation of new words, being applicable both to "pure" phonology and to syntactic organization alike. To this extent, word-formation takes its place in a broad, unified conception of the symbolic system which constitutes a human language.

7. References

Baayen, Harald and Rochelle Lieber
    1991   Productivity and English derivation: A corpus-based study. *Linguistics* 29: 801−843.
Baayen, Harald and Antoinette Renouf
    1996   Chronicling the Times: Productive lexical innovations in an English newspaper. *Language* 72: 69−96.
Bauer, Laurie
    2001   *Morphological Productivity.* Cambridge: Cambridge University Press.
Benczes, Réka
    2007   *Creative Compounding in English. The Semantics of Metaphorical and Metonymical Noun-Noun Combinations.* Amsterdam/Philadelphia: Benjamins.
Bybee, Joan
    1985   *Morphology. A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: Benjamins.
Bybee, Joan
    1995   Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425−455.

Bybee, Joan
  2001   *Phonology and Language Use*. Cambridge: Cambridge University Press.
Dąbrowska, Ewa
  2004   Rules or schemas? Evidence from Polish. *Language and Cognitive Processes* 19: 225−
          271.
Firth, John
  1930   *Speech*. London: Benn's Sixpenny Library.
Gries, Stefan Thomas
  2006   Cognitive determinants of subtractive word formation: A corpus-based perspective. *Cognitive Linguistics* 17: 535−558.
Hay, Jennifer
  2001   Lexical frequency in morphology: Is everything relative? *Linguistics* 39: 1041−1070.
Hilpert, Martin
  2008   The English comparative − language structure and language use. *English Language and Linguistics* 12: 395−417.
Jackendoff, Ray
  2010   *Meaning and the Lexicon. The Parallel Architecture 1975−2010*. Oxford: Oxford University Press.
Kemmer, Suzanne
  2003   Schemas and lexical blends. In: Hubert Cuyckens, Thomas Berg, Rene Dirven and Klaus-Uwe Panther (eds.), *Motivation in Language*, 69−97. Amsterdam/Philadelphia: Benjamins.
Langacker, Ronald W.
  1987   *Foundations of Cognitive Grammar*. Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.
Langacker, Ronald W.
  1991   *Foundations of Cognitive Grammar*. Vol. 2: *Descriptive Application*. Stanford: Stanford University Press.
Langacker, Ronald W.
  2008   *Cognitive Grammar. A Basic Introduction*. Oxford: Oxford University Press.
Mondorf, Britta
  2003   Support for *more*-support. In: Günter Rohdenburg and Britta Mondorf (eds.), *Determinants of Grammatical Variation in English*, 251−304. Berlin: Mouton de Gruyter.
Nesset, Tore
  2008   *Abstract Phonology in a Concrete Model. Cognitive Linguistics and the Morphology-Phonology Interface*. Berlin: Mouton de Gruyter.
Panther, Klaus-Uwe and Günter Radden (eds.)
  2011   *Motivation in Grammar and the Lexicon*. Amsterdam/Philadelphia: Benjamins.
Panther, Klaus-Uwe and Linda Thornburg
  2001   A conceptual analysis of English -er nominals. In: Martin Pütz, Suzanne Niemeier and René Dirven (eds.), *Applied Cognitive Linguistics*. Vol. 2, 149−200. Berlin: Mouton de Gruyter.
Rhodes, Richard and John Lawler
  1981   Athematic metaphors. In: Roberta Hendrick, Carrie Masek and Mary Frances Miller (eds.), *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society*, 318−342. Berkeley, CA: Chicago Linguistic Society.
Saussure, Ferdinand de
  1964 [1916]   *Cours de linguistique générale*. Paris: Payot.
Sinclair, John
  2004   *Trust the Text. Language, Corpus and Discourse*. London: Routledge.

Sosa, Anna Vogel and James MacFarlane
   2002   Evidence for frequency-based constituents in the mental lexicon: Collocations involving
           of. *Brain and Language* 83: 227−236.
Taylor, John R.
   2002   *Cognitive Grammar.* Oxford: Oxford University Press.
Taylor, John R.
   2004   The ecology of constructions. In: Günter Radden and Klaus-Uwe Panther (eds.), *Studies
           in Linguistic Motivation*, 49−73. Berlin: Mouton de Gruyter.
Taylor, John R.
   2012   *The Mental Corpus. How Language is Represented in the Mind.* Oxford: Oxford Univer-
           sity Press.
Tomasello, Michael
   2003   *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge,
           MA: Harvard University Press.
Tuggy, David
   1992   The affix-stem distinction: A Cognitive Grammar analysis of Data from Orizaba Na-
           huatl. *Cognitive Linguistics* 3: 237−300.
Tuggy, David
   2005   Cognitive approach to word formation. In: Pavol Štekauer and Rochelle Lieber (eds.),
           Handbook of Word-Formation, 233−265. Dordrecht: Springer.
Ungerer, Friederich
   2007   Word-formation. In: Dirk Geeraerts and Hubert Cuyckens (eds.), *The Oxford Handbook
           of Cognitive Linguistics*, 650−675. Oxford: Oxford University Press.
van Huyssteen, Gerhard B.
   2010   (Re)defining component structures in morphological constructions: A Cognitive Gram-
           mar perspective. In: Alexander Onysko and Sascha Michel (eds.), *Cognitive Perspec-
           tives on Word Form*ation, 97−126. Berlin/New York: de Gruyter.
Wheeler, Cathy J. and Donald A. Schumsky
   1980   The morpheme boundaries of some English derivational suffixes. *Glossa* 14: 3−34.

John R. Taylor, Christchurch (New Zealand)

11. Word-formation in optimality theory