DE GRUYTER DOI 10.1515/anglia-2013-0066 — Anglia 2013, 131(4): 591–626

Angela Lamberty and Hans-Jörg Schmid

Verbal Compounding in English: A Challenge for Usage-Based Models of Word-Formation?

Abstract: Usage-based models of grammar claim that, in a nutshell, speakers glean their tacit knowledge about language from their linguistic experience in communicative exchanges. Verbal compounding can be considered a challenge to usage-based models: on the one hand, genuine verbal compounding is generally regarded as not being a productive pattern for the formation of new English words; on the other hand, complex lexemes which could very well be the result of such a process, but were created by means of back-formation or conversion, e.g. *to dry-clean, to babysit* or *to house-train*, exist in non-negligible quantities. From a usage-based perspective, then, the question arises as to why speakers of English apparently do not have a productive schema for the creation of genuine verbal compounds at their disposal, even though they are confronted with linguistic input that seems to suggest, at least on the surface, that verbal compounding is indeed a productive process. Given that speakers with no training in linguistics are unlikely to be aware of the formation history of existing verbal compounds, what is the nature of the tacit knowledge they do seem to have that generally prevents them from creating new genuine verbal compounds and from judging them as acceptable when they come across them? This is the question addressed in the present paper. We offer findings from a systematic dictionary-cum-corpus analysis and from an acceptability and comprehension task which strongly suggest that hearers actually do not process novel genuine verbal compounds as compounds, but rely on different processing strategies instead, trying to take recourse to possible base nouns or adjectives and interpreting meanings on the basis of analogies to similar lexical items in the network. A check for the nameworthiness of the concepts denoted by the verbal compounds also seems to be involved. The paper concludes with a set of models representing the processing of different types of genuine verbal compounds and verbal pseudo-compounds, and showing that the ways in which these forms are processed are not conducive to the formation of a productive schema.

Angela Lamberty and Hans-Jörg Schmid, Ludwig-Maximilians-Universität München
E-Mails: angela@lamberty.org, Hans-Joerg.Schmid@anglistik.uni-muenchen.de

1 Introduction

According to Marchand (1969), the English language does not have genuine verbal compounds. Marchand argues that lexemes which superficially look like verbal compounds, such as *to spotlight* or *to stagemanage*, are derived from non-verbal compounds (*spotlight*$_N$, *stagemanager*$_N$) and states quite categorically that genuine "[v]erbal composition does not exist in Present-day English" (Marchand 1969: 100). Complex verbs like the above-mentioned, which were not formed by means of compounding, are referred to as "verbal pseudo-compounds" by Marchand (1969: 101) and divided into two groups, depending on the derivation pattern that underlies their formation. On the one hand, complex verbs can be "derived from a nominal compound (which is almost always a substantive)" (Marchand 1969: 101), mostly as zero-derivations from Noun + Noun or Adjective + Noun compounds. Examples of this type include *to snowball*, *to cold-shoulder* or *to blacklist*. On the other hand, verbal pseudo-compounds can be "derived from [...] synthetic compound[s]" (Marchand 1969: 101), in which case they are back-formations from agent nouns, action nouns or participial adjectives. *To babysit*, for instance, is derived from the agent noun *babysitter*, *to dry-clean* from the action noun *dry-cleaning*, and *to house-train* from the adjective *house-trained*. In contrast to such pseudo-compounds, possible genuine verbal composition would be manifested by lexemes which cannot be traced back to a non-verbal basis, but are actually formed by means of the process of compounding, i.e. by joining two lexemes.

From a rule-based, generative perspective, the claim that the grammar of English does not allow the formation of verbal compounds does not seem to raise serious problems. Two issues might be in need of further explanation, though: firstly, as argued by Adams (1973: 108–109), Kastovsky (1986: 419) and Cho (2002: 161 *et passim*), at least some cases of genuine verbal compounds may in fact exist after all. Secondly, and more significantly, tests that will be reported on here indicate that many language users readily comprehend some invented genuine verbal compounds and rate them as being acceptable. Fabricated hypothetical lexemes like *to *spongeclean*, *to *househop* or *to *fingercomb*, which are not supported by nominal or adjectival bases, turned out to be acceptable, even though the word-formation type as such does not seem to exist. This finding needs to be accounted for from a rule-based perspective.

From a usage-based perspective, however, a much bigger question arises: if the grammatical knowledge of speakers emerges from usage, as is predicted by usage-based models (cf. e.g. Langacker 1988, 2000, Barlow and Kemmer 2000, Bybee 2010), and if usage includes a non-marginal number of complex lexemes which very much look like verbal compounds, are used by speakers and can be

segmented and decomposed by them, then why do speakers apparently not coin genuine verbal compounds? The idea that the existence of verbal pseudo-compounds and of a small number of genuine verbal compounds could spark the production of further genuine verbal compounds and eventually turn verbal compounding into a productive pattern was in fact mentioned prior to the emergence of the label "usage-based" grammar. Kastovsky (1986: 419), for example, referring to Brömser (1985: 111), discusses the possibility that verbal compounding could become a productive process by means of which "new formations are derived on the basis of transformation-like rules". According to him, an increasing number of verbal compounds can actually be analyzed as compounds ("e.g. *to sightsee* = 'see the sights', *to spoonfeed* = 'to feed with a spoon'", 1986: 419) rather than derivations (*to babysit* = 'to act as a babysitter'). However, whether this distinction is of relevance to non-linguists is doubtful, as is indicated by Lieber's claim that "from the point of view of the average native speaker, they [i.e. verbal pseudo-compounds derived by means of back-formation] are just compound verbs" (2009: 361; based on Ackema and Neeleman 2004). If this is true, then why do productive schemas of the type $[A + V]_V$ and $[N + V]_V$ apparently not exist in the minds of language users, even though all the positive evidence that would be required for their existence is available?

Two answers to this question come to mind: firstly, it is possible that speakers of English are in fact able to see behind the structures of apparent verbal compounds after all and – tacitly – realize that they are in fact pseudo-compounds derived from nominal or adjectival compounds. Secondly, language users could be faced with "direct" or "indirect negative evidence" (Chomsky 1981: 8–9) telling them that genuine verbal compounds are not accepted by their fellow-members in the speech community – which, of course, essentially takes us back to the first explanation. The key question, addressed in this paper, is therefore: what is the nature of the knowledge that speakers seem to have about verbal (pseudo-)compounds which seems to prevent them from coining and accepting (most) genuine verbal compounds?

The rationale behind our approach to answering this question is the following: firstly, we investigate the positive evidence that speakers could rely upon when forming verbal compounds. This is done by collecting a large number of existing verbal compounds, most of which are pseudo-compounds, from different sources (see below for more details) and by extracting from this database the morphological, semantic and syntactic characteristics of these lexemes. The point of this methodological step is to identify the regularities which language users can potentially distil in order to form productive schemas, from existing verbal pseudo-compounds (and genuine verbal compounds to the extent that they exist), as predicted by usage-based models. In the second step, in order to

investigate whether these regularities are part of speakers' tacit linguistic knowledge, they are exploited to formulate hypotheses concerning the properties of two types of potentially acceptable verbal compounds, one exhibiting the typical characteristics, and one violating them. The validity of these hypotheses is tested in a comprehension and acceptability study. The results of this study suggest that speakers are acutely aware not only of the typical characteristics of verbal pseudo-compounds but also of the limits of the productivity of genuine verbal compounding. Speakers do seem to have both positive evidence – of what is out there – and negative evidence – of what could be out there, but is not – at their disposal. Finally, on the basis of the results of this study, a model is deduced which captures the way in which novel verbal compounds and pseudo-compounds are processed. This model captures the reasons why a productive schema for genuine verbal compounds does not seem to be available to speakers of English even though the input for such a schema appears to be.

Before we embark on a report of our study, we will briefly recapitulate some of the key publications on verbal compounds and pseudo-compounds.

2 Previous Work on Verbal Compounding

Researchers do not agree on whether or not (genuine) verbal compounding is a productive word-formation type. As we have seen, Marchand (1969) was certain that it is not. Adams (1973: 108–109) regards verbs giving cleaning instructions such as *to handwash* or *to coldrinse*, which are typically found in imperative forms on product labels, as examples of genuine verbal compounds. Moreover, she gives examples such as *to chain-drink*, an analogy to the verb *chain-smoke* (Pennanen 1966: 115), or nonce-formations such as *to cathedral-look* (Adams 1973: 108). Brömser (1985: 111) and Kastovsky (1986: 419) follow Adams's lead. In a later publication, however, Adams (2001: 100–109) devotes a full chapter to "verb compounds" and makes it very clear that the vast majority of verbal compounds are not genuine verbal compounds but derivations. In contrast, Cho (2002: 242) claims that non-derived verbal compounds exist in considerably larger numbers than assumed by Marchand (1969: 100), Adams (1973: 109) and others and argues that verbal compounding has been, at least to some extent, a productive word-formation type in English.

Why do genuine English verbal compounds not exist or seem to be extremely rare? One answer to this question may well lie in the relatively fixed English word order, deviations from which result in unacceptable constructions. According to Pennanen (1966: 111), constructions in which the direct object pre-

cedes the verb, as in *to *meat-eat* instead of *to eat meat*, are generally judged ungrammatical by speakers of English. This could be one reason why the English language does not allow such Object + Verb compounds. However, since in German, which is more flexible regarding its word order and allows objects to precede the verb, comparable verbal compounds such as **fleischessen* or **hauskaufen* do not exist either, fixed word order alone does not seem to be a sufficient reason for the non-existence of verbal compounds. Furthermore, the fact that syntactic word order is often reversed in other, productive types of compounds (*He plays tennis → tennis player*) indicates that the word order violation is too weak an argument.

The basic structural and typological characteristics of English also seem to forbid incorporating constructions, in which "a noun stem is compounded with a verb stem" (Mithun 1986: 32), which can be found in certain polysynthetic languages. This issue has been addressed in a number of publications, particularly within the framework of Functional Grammar. Several authors have tried to test the limits of regarding English verbal (pseudo-)compounds as instances of noun incorporation. Hall (1956: 83–88), for example, noted a number of parallels between English verbal compounds and noun incorporations in polysynthetic languages. Kirchner (1959: 302) even claimed that a tendency towards a "new synthesis" could be observed to work against the otherwise increasingly analytical structure of English. Brömser (1985: section 3.3), too, suggested that verbal compounding could develop into a productive noun incorporation pattern in English. But even though English verbal compounds do share some formal features with noun incorporations in polysynthetic languages, they do not show the characteristics observable in typical incorporating constructions and are thus better not regarded as instances of noun incorporations. The most substantial counterargument is the one mentioned in the introduction, viz. that verbal compounds derive from nominal compounds rather than being truly compounded. Genuine incorporating constructions, in languages that feature them, also provide systematic syntactic and information-dynamic options that are not available in English verbal pseudo-compounds. As Mithun (1985: 371–373) demonstrates, incorporating constructions allow the speaker to control the focus of attention by alternating between incorporations and free nouns. Accordingly, new pieces of information that are supposed to be highlighted are preferably realized as independent nouns, whereas information that does not require special attention can be incorporated and thus backgrounded in discourse. In contrast to such constructions, in which it is mostly a noun that is incorporated, English verbal pseudo-compounds are not restricted to Noun + Verb combinations, as lexemes such as *to dry-clean* (A+V) or *to drip-dry* (V+V) show. Finally, Mithun (1984: 847–848) observes that noun incorporations always possess a se-

mantically equivalent syntactic counterpart. Thus, an incorporating structure that can be translated as 'I-reindeer-slaughter' implies the existence of an equivalent syntactic phrase 'I-slaughter-reindeer'. This aspect leads to problems when comparing noun incorporation to verbal compounding in English: not only metaphorical compounds such as *to cherry-pick* ('to choose the best parts of something') but also many other verbal pseudo-compounds do not have direct syntactic equivalents precisely because they are not compounded but derived; cf. the notorious case of *to babysit*, which is more appropriately paraphrased by 'act as a babysitter' than 'sit by the baby'.

This observation is the point of departure for Ackema and Neeleman's (2004) so-called 'Morphosyntactic Competition Theory', which suggests that verbal compounding in English is 'blocked' by syntax. The authors argue that syntax and morphology are independent systems which compete for the right to combine elements into complex constructions. As English has been shifting towards the analytical rather than synthetic pole and thus prefers syntactic phrases over morphological ones (Ackema and Neeleman 2004: 41), what would be the point of combining a noun and a verb to form a verbal compound if a sentence expressing the same content is an entrenched and conventional option? Morphological combination would thus only be asked for in cases where syntactic merger cannot sufficiently express the semantic relation that holds between the constituents, i.e. when there is no competition between the two options. The Morphosyntactic Competition Theory thus predicts that hypothetical verbs like *to *meat-eat* or *to *house-buy* are excluded from the pool of possible combinations, since their meaning could be paraphrased as 'to eat meat' and 'to buy a house' respectively, thus containing no additional elements that would distinguish them from the syntactic phrase. However, a lexeme like *to colour-code* cannot be paraphrased as 'to code colours', but means 'to code with colours'. The first constituent has the role of an adverbial in the underlying sentence, not that of an object. This means that different categories merge in the morphological combination (noun and verb) and in the corresponding syntactic option (verb and prepositional phrase). As a result the syntactic competitor does not block the complex verb according to Ackema and Neeleman (2004: 60), which explains the existence of the verbal compound. The problem, however, is that *to colour-code* is not a genuine verbal compound but a derived verbal pseudo-compound, and the same goes for the vast majority of complex verbs that include non-argumental relations as constituents. While Ackema and Neeleman's theory thus correctly excludes as impossible all those combinations in which the first element serves as an argument of the verb, it does not explain why genuine verbal compounds incorporating non-argumental relations do not exist either. In addition, further empirical evidence casts doubt on their broad

generalization, as some INSTRUMENT relations exist in certain verbs, for example in *to handfeed* or *to smokesignal*, whereas the same relation does not work in others, cf. *to *hand-eat* or *to *crutchwalk*.

In his unpublished *Habilitationsschrift* on composite verbs from a synchronic and diachronic perspective, Cho (2002) claims:

> Die Auswertung des *OED* 2 hat gezeigt, dass in der Geschichte des Englischen Verbkomposita nicht durchgängig abgeleitet sind. Für alle Perioden des Englischen finden sich sowohl abgeleitete wie nicht-abgeleitete. Dabei überwiegen von den Zahlen die ersteren. Auf sie entfallen nach den Belegen im *OED* 2 etwa Dreiviertel der Verbkomposita. Das verbleibende Viertel wird von den nicht-abgeleiteten gestellt. (Cho 2002: 242)
>
> [The analysis of *OED* 2 has shown that in the history of English not all verbal compounds are derived. For all historical periods of English both derived and non-derived verbal compounds can be found, with the former outweighing the latter in quantitative terms. Derived verbal compounds account for approximately three quarters of the attestations in *OED* 2. The remaining quarter is contributed by the non-derived ones.]

What the author also shows, however – and this will be supported by the present study –, is that a large number of genuine verbal compounds are formed on the basis of constructional schemas derived from existing verbal pseudo-compounds or, more rarely, potential genuine verbal compounds (2002: ch. 10) essentially by analogy (see below). This limitation casts doubt on Cho's general conclusion (2002: 161) that verbal compounding has been and still is a productive word formation type in English – a doubt which is confirmed by our findings and by Cho's (2002: ch. 11) own analyses of usage frequencies in corpora, which indicate that the majority of complex verbs do not manage to diffuse and acquire an even moderate degree of everyday currency. Even if it happens that for some reason or other genuine verbal compounds are coined by individual speakers (or rather writers), these formations do not seem to be accepted and used by other members of the speech community and therefore do not manage to become a permanent part of the lexicon.

In their use in actual speech situations, verbal pseudo-compounds are subject to considerable limitations concerning the dispersion of their word-forms (cf. Cho 2002: 182 *et passim*) and to uncertainties about the past tense forms of irregular verbs (Adams 2001: 102–103) – cf. e.g. *babysat* or *joyrode*. Extreme cases can in fact be claimed to have more or less defective paradigms. Speakers using verbal pseudo-compounds definitely favour their base-forms and *-ing*-forms, which may indeed show the strong association of the complex verbs to their nominal bases, many of which are synthetic formations ending in *-ing*. To give some examples, in the *BNC*, which reflects the state of English as of the late 1980s and early 1990s, not a single form other than base and *-ing* was found for the verbs *proofread* and *lipread*. For *gatecrash*, the *BNC* does have as

many as 11 past tense forms out of a total of 31 tokens of the lexeme. *Babysit*, arguably one of the most frequently used verbal pseudo-compounds, has no more than 3 out of 155 occurrences with a third-person *-s*; in addition, there are 5 occurrences of the past tense form *babysat*, but all others are base or *-ing*. It is not unlikely that the existence of these strongly skewed paradigms is a symptom of the special status of these verbs and serves as a covert indicator to language users that something "is wrong" with them.[1]

From a cognitive-linguistic perspective, which will inform the present paper, the comparison of competing syntactic and morphological constructions mentioned above promises to benefit from integrating the concept of hypostatization (Leisi 1975: 26, Leech 1981: 32, Lipka 1977: 161, Schmid 2008: 5–9). This notion describes the phenomenon that "the existence of a particular word creates the impression that there is a corresponding thing or entity to which the word refers" (Schmid 2008: 5). In contrast to a corresponding syntactic construction, e.g. *to code with colours*, the use of one word, *to colour-code*, suggests the existence of an accepted practice denoted by that word, which has the quality of a shared social gestalt (cf. Mithun 1984: 848, see below). Hypostatization is a pervasive phenomenon which concerns simple as well as complex lexemes; however, the hypostatizing potential, i.e. the "'concept-forming' power of the word", as Leech (1974: 37) calls it, is stronger for some word classes than for others. Nouns, which suggest the existence of a temporally stable cognitive category of thing-like entities, possess a higher hypostatizing potential than adjectives and verbs and are thus more likely to be coined in order to name new concepts (Schmid 2008: 6–9). As already suggested by Grimm (1877: 577), it could be that the nonexistence of genuine verbal compounds is grounded in the fact that nouns, which prototypically refer to persistent, unchanging states of affairs, radically oppose the dynamic nature inherent in verbs.[2] To this we could

1 In the more recent *Corpus of Contemporary American English* (*COCA*, Davies 2008–), out of a random selection of 100 attestations of the form *proofread*, 10 were past tense and 14 past participle forms. The 3rd person singular form *proofreads* is attested 5 times in the whole 450 million-word corpus. The past tense form *babysat* is attested 94 times out of 421 occurrences of the lexeme *babysit* in *COCA*, the form *babysits* 23 times. Overall, the comparison with the *BNC* indicates that the reluctance to use inflected forms decreases as frequent verbal (pseudo-)compounds become more established and are accepted as 'normal' verbs by speakers.

2 "Sein ganzes wesen ist thätigkeit, entgegengesetzt der ruhe des nomens. Bei dem nomen soll eben die composition bleibende zustände im ausdruck feßeln. Das verbum, nach zeit und modus regsam und bewegt, übt einen viel zu manigfaltigen einfluß auf das nomen aus, als daß er nicht durch zusammensetzungen sollte gehemmt werden" (Grimm 1877: 577). [Its whole essence is activity, in contrast to the stillness of the noun. With the noun, the role of the composition is to encapsulate unchanging states in the expression. The verb, active and in motion

add the complementary observation that the dynamic and relational concepts typically encoded by verbs do not lend themselves to the kind of head-modifier relationship found in prototypical determinative nominal and adjectival compounds. While these arguments are weakened by the indisputable existence and frequent use of verbal pseudo-compounds such as *to babysit*, *to table-hop* or *to headhunt*, it can still be assumed that verbal compounding is a less 'natural' conceptual process than adjectival and especially nominal compounding and that speakers therefore need additional motivations to coin and accept verbal compounds.

In sum, while previous work in the field has provided a number of important insights and opened up avenues for the investigation of genuine verbal compounds and verbal pseudo-compounds, it falls short of answering the question we are addressing in this paper, i.e. the question of whether speakers of English have specific knowledge about the characteristics of verbal pseudo-compounds at their disposal which prevents them from coining and accepting genuine verbal compounds. In order to get a detailed picture of the positive evidence potentially available to speakers of English, we will now look at the morphological, semantic and syntactic characteristics of established verbal pseudo-compounds and, to the extent that they exist, genuine verbal compounds.

3 The Dictionary-cum-Corpus Study: The Characteristics of Existing Verbal Pseudo-Compounds and Genuine Verbal Compounds

3.1 Material

For a more profound and systematic analysis of established verbal pseudo-compounds, lexemes from two different sources were analyzed: the *Longman Dictionary of Contemporary English* (*LDOCE*) (CD-ROM version 2005) and the extensive appendix of Cho's (2002) study on verbal compounds mentioned in section 2 above. With regard to the *LDOCE*, all compound verbs recorded there were extracted manually and filtered and classified according to certain criteria. To ensure comparability with genuine verbal compounds, only complex verbs whose final constituent was verbal were selected, thus excluding

according to tense and mode, exerts such a multifarious influence on the noun that it should not be impeded by means of compounding.]

$[N+N]_V$ compounds such as *to pigeonhole* or *to bootleg*, as they cannot be formed by a genuine verbal compounding process. This method was also applied to the second source, Cho's appendix. This appendix consists of 38 pages containing verbal (pseudo-)compounds gathered from three corpora (*Broadcast News*, *Berliner Korpus* and *Time Korpus*) and four monolingual dictionaries (*The American Heritage Dictionary of the English Language*, 3rd ed., 1992; *Merriam-Webster's Collegiate Dictionary*, 10th ed., 1996; *Webster's New World College Dictionary*, 3rd ed., 1996; and *The Newbury House Dictionary of American English*, 1999). Folk-etymologies (e.g. *to piggyback*, originally a combination of *to pick* and *pack*, *OED* s.v. *piggyback*, v.), reduplications (e.g. *to flip-flop*, *to see-saw*), neoclassical compounds (e.g. *to stereotype*) and blendings (e.g. *to guesstimate*, *to breathalyze*) were eliminated on the grounds that their constituents cannot be used as free lexemes and are thus not proper cases of composition. A total of 627 lexemes remained as relevant material for the analysis.

Although all these verbs are interesting for the purpose of this study, it should be kept in mind that a considerable number of the items listed by Cho (2002) are highly marked, used very rarely or may even be regarded as nonce-formations. For example, Google searches (carried out on 14 August 2013) have produced no verbal attestations of the forms *to broadstroke*, *to carshop* and *to handsnap* mentioned by Cho, which casts doubt on their existence as institutionalized and entrenched lexemes.

3.2 Analytical Categories

Inspired by Lipka's (1983) multi-level approach to word-formation – which was extended by Schmid (2011) – and relying on further sources (see below), the material was analyzed with regard to the following criteria:

- the morphological shape of the verbs
- the temporal structure of the complex lexeme
- the semantic relation holding between the constituents
- the existence of word-families, and
- the nameworthiness of the concepts denoted by the verbs

More details concerning these criteria (see also Lamberty 2012) will be given in the next section, which will summarize the insights gained from the dictionary-cum-corpus study.

3.3 Results of the Dictionary-cum-Corpus Study

3.3.1 Morphological Shape

The first criterion focuses on the formal characteristics of the lexemes and takes into account their constituent morphemes, their word classes and functions in the compound. With regard to morphological shape, there is a strong preponderance of Noun + Verb combinations (e.g. *to handcraft*, *to skywrite*; 69.2%), followed by Adjective + Verb (*to softland*; 22.0%) and Verb + Verb (*to stirfry*; 8.8%) combinations. Our database of 627 lexemes contains only a single element which includes more than two lexical morphemes, viz. the verb *to fine-toothcomb*. All others consist of two constituents only.

3.3.2 Temporal Structure

On the level of temporal structure, the situation type or *aktionsart* is considered on the basis of Vendler's (1957) taxonomy of verbs. More than half of the verbs in the database, 56%, are ACTIVITY verbs (e.g. *to fundraise, to headhunt*). ACCOMPLISHMENT verbs (e.g. *to ringfence, to spindry*) account for exactly one third of the material, while ACHIEVEMENT verbs (e.g. *to namedrop, to skyrocket*; 9%)[3] and STATE verbs (e.g. *to daydream, to lobbysit*; 2%) are much rarer. This indicates that the overwhelming majority of verbal (pseudo-)compounds are dynamic verbs.

3.3.3 Semantic Relations

The next level of analysis deals with the internal semantic relations between the elements of the complex lexeme. These relations are classified according to a modified selection of Fillmore's (1968) account of semantic roles. In our context, the distinction between participant roles and circumstantial roles will turn out to be of key importance. Participant roles are those filling an obligatory slot in the valency structure of the verb functioning as second constituent of the verbal (pseudo-)compound, i.e. AGENT, PATIENT and THEME. Circumstantial roles,

3 In our operationalization of Vendler's categories, ACHIEVEMENTS differ from ACCOMPLISHMENTS in that the former have the feature [–DURATIVE], while the latter are [+DURATIVE]. In addition, ACCOMPLISHMENTS are clearly [+TELIC], while this feature is optional for ACHIEVEMENTS.

in contrast, are typically realized as optional adverbials; they are classified as INSTRUMENT, MANNER, CAUSE, TIME or PLACE.

The most striking finding concerns the fundamental difference between participant and circumstantial roles. The participant roles of AGENT (e.g. *to teamteach*), PATIENT (e.g. *to tongue-tie, to gift-wrap)* and THEME (e.g. *to headhunt, to cherry-pick*), which would typically incorporate the subject or object of the underlying sentence, in fact occur very rarely in verbal compounding. Only every seventh case (14%) is of this type. What is more, while such verbs at first sight seem to contain one of these participant roles, a closer look reveals that the first constituents resemble the direct object/subject of the verb only superficially. From a syntactic perspective, in spite of the incorporation of an apparent object, these complex lexemes are mostly still transitive and require an object (e.g. *to giftwrap a book*). For instance, while the first constituent of *to giftwrap*, *gift*, seems to play the role of PATIENT ('to wrap a gift'), it actually represents the MANNER component ('to wrap sth like a gift') in the underlying sentence. From a semantic perspective, it is remarkable that many of these compounds have figurative meanings, which usually rest on the first constituent. For example, *cherry* in *cherry-pick* stands metaphorically for the best part, *head* in *headhunt* metonymically for a person. All this strongly suggests that true, rather than superficial, AGENT/PATIENT/THEME + Verb compounds are predominantly figurative, which excludes lexemes like *to *meat-eat* 'to eat meat', *to *housebuy* 'to buy a house' or even less likely combinations such as *to *babycry*, which would have to be derived from 'the baby cries' if it was an AGENT-type. A related idea can also be found in Cho (2002: 77), who points out that direct objects cannot be included in verbal compounds.

Circumstantial roles, in contrast, lend themselves very well to the formation of verbal compounds. The vast majority of 84% of the verbs collected in the database feature the roles of PLACE (*to skywrite* 'to write in the sky'; 35% of all circumstantial roles), MANNER (*to rough-handle* 'to handle in a rough manner'; 27%), INSTRUMENT (*to handpaint* 'to paint with the hand'; 24%), TIME (*to spring-clean* 'to clean in the spring'; 7%) or CAUSE (*to joyride* 'to ride for sheer joy'; 7%). In short, the encoding of circumstantial roles mapped onto non-obligatory sentence constituents clearly dominates in verbal (pseudo-)compounds over that of participant roles encoded by core constituents.

3.3.4 Word-Family Effects

Most complex lexemes in our database are not isolated islands but densely connected with other complex lexemes by formal and semantic relations. These interconnections are established by both the first and the second constituent (cf.

Cho 2002: ch. 10). The verb *to **hand**write*, for example, is linked to several other lexemes in the database via its first constituent *hand-*. In total, 21 lexemes of this kind are attested in the database, among them *to **hand**pick*, *to **hand**feed*, *to **hand**wash* and *to **hand**stamp*. The second constituent *-write* is present in the verbs *to ghost**write***, *to sky**write*** and *to type**write***. Interconnections of this type can be observed for the vast majority of the lexemes in our database (e.g. *to jobhunt* with *to **job**share*, *to **job**hop*, *to fox**hunt***, *to head**hunt***, *to bargain**hunt***). As illustrated by figure 1 and also noted by Cho (2002: ch. 10), these relations conspire to form a densely knit network based on formal and semantic similarities. It is likely that this network contributes to the triggering of analogy-based formations, as suggested, among others, by Adams (1973: 108), Kastovsky (1986: 419), Hansen et al. (1990: 63, 136–137) and Cho (2002: 125–128). Whether these word-family and family-size effects (cf. Schreuder and Baayen 1997, de Jong et al. 2002, Booij 2005, de Vaan, Schreuder & Baayen 2007) actually increase the acceptability of new verbal complex lexemes and potentially facil-

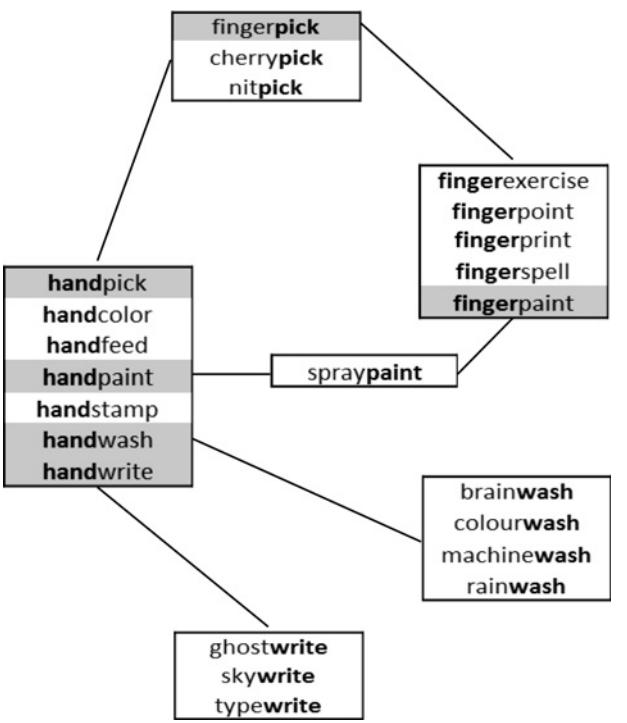

Fig. 1: Illustration of network of word-families based on formal and semantic similarities

itate their formation was tested in the questionnaire study reported on in section 4.

3.3.5 Nameworthiness

The final criterion used in the classification of the material concerned the issue of nameworthiness, i.e. the question of whether there is a referent category that is in need, or at least worthy, of being named and thus hypostatized as a socially relevant conceptual gestalt (Zimmer 1971/1981: 249,[4] Downing 1977: 837–838, Mithun 1984: 848, 1985: 365–366). While nameworthiness is undoubtedly a prime motivation for coining, accepting and using (novel) complex lexemes, the phenomenon is notoriously difficult to operationalize (cf. Kerremans 2012: section 3.3.3). Therefore, for the purposes of the questionnaire study, two striking observations emerging from a close semantic analysis of the verbs in the database were focused on: metaphorical and metonymical meanings, on the one hand, and meanings that expressed noticeable deviations from a normal, expectable procedure, on the other.

As to the first criterion, the number of lexemes with figurative elements was very conspicuous. Verbs like *to sweet-talk*, *to frostbite* or *to cradle-rob* are only a few of many cases that could be listed here. Whereas in *to sweet-talk* the first constituent is used metaphorically (*sweet* in the sense of 'pleasing to the ear'), it is the second one in *to frostbite* (with *bite* conceptualizing the pain that can be caused by intense cold as the pain that follows the bite of a fierce animal). In the last example, *to cradle-rob* 'to be/fall in love with a much younger person', a whole scene is evoked, which both metaphorically and metonymically relates to the scenario described.

The second characteristic that seems to account for the nameworthiness of a large number of verbal (pseudo-)compounds will be referred to as 'deviation from the norm' here. This describes the fact that a strikingly high number of corpus verbs derive their relevance from a particularly uncommon or particularly noteworthy manner of action. They denote activities that are in evident contrast to a 'normal' procedure. The lexeme *to vacuumclean*, for instance, describes a way of cleaning that is deviant from a normal cleaning process in that it employs a (then) novel method. Further examples include *to forceland* ('to make a forced landing', *OED* s.v. *forced*, adj.), *to speed-dial* or *to waterski*.

4 Zimmer (1971/1981: 249) also talks about the "classificatory relevance" of concepts.

3.3.6 Summary of the Dictionary-cum-Corpus Study

The analysis of the material collected in the database has yielded the following insights: prototypical verbal pseudo-compounds

- consist of two morphemes only
- have a noun as their first and a verb as their second constituent
- denote ACTIVITIES or ACCOMPLISHMENTS
- profile circumstantial, rather than participant relations to the verbal head, most typically those of PLACE, MANNER and INSTRUMENT
- belong to tightly-knit and sizable word-families
- are nameworthy in that they either denote some sort of deviation from a norm or have figurative meanings

4 The Questionnaire Study: Investigating the Acceptability and Comprehensibility of Novel Verbal Pseudo-Compounds and Genuine Verbal Compounds

The observed characteristics of typical verbal pseudo-compounds can now be transformed into hypotheses relating to what could be referred to as 'potentially genuine' verbal compounds. Roughly speaking, the usage-based framework would suggest that the properties of typical pseudo-compounds are more likely to serve as input for a potential schema of verbal compounds than the properties of very unusual verbal pseudo-compounds. If it is further assumed that the availability of such a schema would have an effect on whether or not and to what degree a given novel compound is acceptable and comprehensible, then it can be expected that the acceptability and comprehension of invented verbal compounds by native speakers varies depending on whether or not they meet the conditions that were found to be dominant in the corpus. Since it was impossible to target all six variables listed in section 3.3.6, the findings concerning the morphological make-up of the items in the database and their temporal structure were not implemented as target variables in the test design, but neutralized in such a way that only test stimuli corresponding to the most frequent types were constructed. This means that all test stimuli were constructed as N+V compounds and denoted ACTIVITIES or ACCOMPLISHMENTS. In this way, it was

avoided that these two variables would confound the results in an uncontrolled manner.[5]

4.1 Hypotheses

The first hypothesis, tested by means of an acceptability rating task, addresses the fundamental difference between genuine verbal compounds and verbal pseudo-compounds:

> Hypothesis 1: Other things being equal, novel genuine verbal compounds are judged as being less acceptable than novel verbal pseudo-compounds.

Thus, hypothetical lexemes of the type *to *palm-read*, which is a derivation from the agent/action nouns *palm-reader/palm-reading*, are expected to yield higher acceptability rates than lexemes like *to *fingercomb*, where no derivation base exists.

The second hypothesis, tested by means of an acceptability rating task and a comprehension task, is derived from the finding that the items in the database are densely interrelated by morphological and semantic similarities of first or second constituents:

> Hypothesis 2: Other things being equal, word-families have a positive effect on the comprehension and acceptability of novel verbal compounds.

Accordingly, a hypothetical lexeme such as *to *hand-signal* should be readily accepted and understood on the grounds that it is embedded in a dense network of related lexemes, whereas the absence of such word-families is supposed to have a negative impact on the comprehension and acceptability of novel verbal compounds.

The third hypothesis concerns the concept of nameworthiness.

> Hypothesis 3: Verbal compounds which denote deviations from a norm and verbal compounds with figurative meanings yield a higher acceptability and are more readily understood than verbal compounds whose concepts do not display a particular nameworthiness.

Thus, verbs like *to *speed-date,* which denotes a novel, unusual way of dating, or *to *clod-hop* (from the noun *clod-hopper* 'clumsy person'), which is metaphorical, are expected to yield better results in both the acceptability and com-

5 A small number of invented stimuli (e.g. *to *stamp-collect, to *rumour-spread* and *to *window-clean*) included participant roles as constituents to check how informants would react to them. This variable was not targeted in the analysis, however.

prehension task than verbs like *to *hand-signal* or *to *figure-skate*, which do not seem to be particularly nameworthy.

4.2 Stimuli

In order to test the three hypotheses, lexemes of three different types were used as test stimuli:

- invented genuine verbal compounds, i.e. verbs like *to *spongeclean*, where no possible derivation base exists
- invented verbal pseudo-compounds, i.e. verbs like *to *palm-read*, which have been derived from an established nonverbal basis, in this case *palm-reader*
- established pseudo-compounds from the database used as distractors and in order to have a base-line to which to compare the results for the invented stimuli

In total, the set of test lexemes comprised 74 verbs: 35 genuine verbal compounds, 36 verbal pseudo-compounds and three distractors (see the appendix for a full list of the stimuli used and the results obtained). The lexemes were invented in such a way that they differed systematically with regard to the variables word-family and nameworthiness, covering the whole range from verbs displaying no word-families (e.g. *to *weed-sow*, *to *fabric-soften*) to large word-families (e.g. *to *househop* related to *to housebreak, to houseclean, to house-sit* and *to barhop, to tablehop, to jobhop*) and from not nameworthy (e.g. *to *table-eat, to *stamp-collect*) to highly nameworthy lexemes (e.g. *to *fingercomb, to *palm-read*). As pointed out above, nameworthiness was operationalized by the presence of figurative elements and the existence of meaning components related to deviations from the norm.

4.3 Tasks, Procedure and Participants

The stimuli were tested on comprehension and acceptability by means of an online questionnaire. Due to the high number of 74 test verbs, the questionnaire was split into three sub-questionnaires each containing 24 or 25 items while ensuring that each test lexeme appeared only once per questionnaire. 108 native speakers of English agreed to complete one of these three-part questionnaires anonymously. Among the 108 participants were 58 speakers of American English, 46 speakers of British English and 4 speakers of Australian English. On average, they were 38.5 years old, the youngest participant being 17, the eldest 82 years old. In the first part, the comprehension test, participants were con-

fronted with the base form of a novel test lexeme (e.g. *to *lion-tame*) for which they were asked to choose between two pre-formulated meaning options (for *lion-tame*: 1) 'to reduce the fierceness of a lion and render it docile' and 2) 'to control and calm down exuberant kids or pupils'). These two options differed as to whether or not they contained a metaphor and whether or not their meaning (for fictitious pseudo-compounds) was derived from the nominal basis. Additionally, participants were also given the possibility to insert a completely different semantic paraphrase if none of the alternatives seemed plausible to them.

The second part of the questionnaire was designed to test the acceptability of the test items. Participants had to rate each lexeme on a four-point scale ranging from −2 ("This word sounds completely unacceptable") to +2 ("This word sounds acceptable/I could imagine it being used").

In the third part of the questionnaire, participants were asked to provide some personal data such as age, nationality and native language. This information was used mainly to identify non-native speakers, whose data were not used.

The statistical analysis of the collected data was performed by means of an ordinal regression using the logit function. Since the target variable in the acceptability rating is an ordinal scale, this is the appropriate model. The advantage of a regression is that it allows for analyses of the estimated effect sizes of individual predictors such as the presence of base nouns or word-families while taking into account the effects of the other variables currently not focused on. The logit is the natural logarithm of the odds of the outcome of an event (i.e. a dependent variable) given a predictor variable. While logit scores are maximally useful and precise from a statistical point of view, they defy intuitive interpretations; therefore, we will also render the more accessible transformed odds ratio, which, in the present case, essentially indicates the odds (= probability divided by converse probability) with which a certain group of test items are accepted compared to a reference group under the circumstances captured as an independent variable.[6] As far as the results for individual stimuli are concerned, we will provide two types of information: firstly, the median of the ratings of the participants who rated an item, which is the mathematically appropriate measure, and secondly, as the median provides only very general information which does not reveal subtler differences between individual items, the arithmetic mean of the ratings of the participants.

6 We would like to thank the Statistical Consulting Unit at the Institute for Statistics of Ludwig-Maximilians-Universität, Munich, for statistical consulting, as well as its director, Helmut Küchenhoff, for advice on the presentation of the results.

4.4 Results

4.4.1 Hypothesis 1

The first hypothesis focuses on the difference between genuine and derived verbal compounds, the former of which were expected to yield lower rates in the acceptability task than the latter. Indeed, as the results demonstrate, verbs such as *to *knife-open* or *to *crutchwalk*, which are not supported by existing lexemes, overall receive more negative ratings than derived verbs like *to *lion-tame* (*lion-tamer*$_N$) or *to *figure-skate* (*figure-skating*$_N$). As a consequence, the category of genuine verbal compounds (GVC) displays negative effects in comparison to the category of verbal pseudo-compounds (VPC).

Parameter	Estimate	Estimated odds ratio
[GVC]	–1.354 ($p < 0.001$)	0.26
[VPC]	0	

The logit coefficient for GVCs, compared to VPCs, is –1.354, the transformed estimated odds ratio 0.26. This means that the odds for a given lexeme to be rated as being acceptable are reduced by a factor of 0.26 for genuine verbal compounds (GCV) as opposed to verbal pseudo-compounds (VPC), when all other factors are fixed.

The comprehension task of the questionnaire, in which participants could choose between two alternating meaning paraphrases, also revealed highly interesting tendencies with regard to the first hypothesis. For one thing, the rate of participants who refused to answer this part is considerably higher for genuine verbal compounds than for verbal pseudo-compounds. Whereas only 10% of the participants regarded neither option as a plausible paraphrase for the class of verbal pseudo-compounds, a total of 24% abstained from choosing one of the alternatives offered for genuine verbal compounds. For derived verbs, one meaning alternative was always formulated in accordance with the meaning of the base lexeme. This option was preferred over the fabricated one without exception. For example, the verb *to *foot-drag* was understood in the same way as the underlying action noun *foot-dragging*, i.e. as 'to deliberately delay something or be slow to do something'.

4.4.2 Hypothesis 2

The next hypothesis concerns the presence of word-families, which was supposed to facilitate the acceptability and comprehension of novel lexemes. As

indicated in the table below, novel verbs which are embedded in word-families clearly benefit from this network:

Parameter	Estimate	Estimated odds ratio
[–WFE]	–0.454 ($p < 0.001$)	0.64
	[+WFE]	0

The coefficient measuring the effects of an absence of word-families (–WFE) is negative (–0.454) and highly significant ($p < 0.001$). The corresponding odds ratio of 0.64 indicates that the odds that a novel verb which is not formally and semantically related to already established ones (–WFE) is accepted is 0.64 times lower than the odds for an otherwise identical verb exhibiting word family effects (+WFE). *To *household*, for instance, benefits from an extensive network of established verbs including *to jobhop, to tablehop* or *to barhop* and was well accepted.

With regard to comprehension, an interesting observation concerned the fact that for verbs related to word-families, the semantic relation underlying the majority of lexemes in the network was adopted. For example, the element *hand-* in established verbs like *handpick* or *handstamp* mostly refers to the INSTRUMENT. When offered the two paraphrases a) 'to eat without using cutlery, to eat with bare hands' (INSTRUMENT) and b) 'of tame animals: to eat straight of a person's hand' (PLACE), 84% of the test participants opted for option a) encoding the INSTRUMENT relation also found in the established verbs. This tendency was more pronounced for genuine verbal compounds than for verbal pseudo-compounds, since the semantics of the base lexeme overruled word-family effects in the latter.

4.4.3 Hypothesis 3

The last hypothesis is concerned with the presence of nameworthiness and assumes that figurative meanings and norm deviations enhance both the acceptability and the comprehensibility of novel verbal compounds. The figures in the table below display the coefficients for non-nameworthy concepts (–NW), those containing a metaphor (+NW (metaphor)) with reference to those indicating a deviation from the normal procedure (+NW (norm deviation)).

Parameter	Estimate	Estimated odds ratio
[–NW]	–0.212 ($p < 0.001$)	0.81
[+NW (metaphor)]	–0.736 ($p < 0.001$)	0.48
[+NW (norm deviation)]	0	

As can be seen, norm deviations (+NW) yield the best results in the acceptability rating task. Verbs encoding unexpected, non-canonical ways of doing things such as *to *speed-date* or *to *mudbathe* exhibit higher odds of being accepted than non-nameworthy lexemes, because the striking circumstances of the activity highlighted by the verbs seem to assign some sort of raison d'être to them. In contrast, novel verbal compounds for concepts without any kind of nameworthiness (–NW) display negative effects of –0.212 (estimated odds ratio 0.81). Quite astonishingly, however, and against what was expected, the logit coefficient for metaphorically enriched verbs, –0.736, is even stronger. Essentially, this means that the presence of metaphorical components severely reduces the chances of a novel lexeme to be considered acceptable. Both genuine verbal compounds like *to *weed-sow* (fabricated to mean 'to put out a rumour, which subsequently spreads with immense rapidity') and verbal pseudo-compounds like *to *curtain-raise* ('to perform as an opening band for the main act' from *curtain-raiser*$_{N}$) yielded negative results in the acceptability test.

While this seems to indicate that metaphorical meanings impede the potential "success" of verbal compounds, the comprehension tests do not entirely support this conclusion.[7] In this test a metaphorical meaning paraphrase was systematically tested against a literal one. When presented with two such predefined options, a strong tendency in favour of the figurative alternative could be observed, which was preferred over the literal meaning paraphrase in 63% of answers for genuine verbal compounds. An analysis of the meaning paraphrases provided by the participants revealed two further interesting points which partly support this interpretation. Firstly, the participants readily make use of metaphorical language when asked to imagine a plausible meaning. Secondly, the associations differ crucially for novel verbal pseudo-compounds and novel genuine verbal compounds. The former, since they are based on established nominal or adjectival combinations, trigger meanings in which the verb as a whole is metaphorized. For example, one participant offered the paraphrase 'to suggest a solution to a problem that looks/sounds good but is ineffective' for the verb *to *figure-skate*. In contrast, genuine verbal compounds cannot rely on a compound basis; as a consequence, the paraphrases triggered by such verbs are not derived from the combination as a whole, but rather from one of its constituents. *To *timecut*, for instance, evoked the lexeme *shortcut*

7 A reviewer of this paper has rightly pointed out that the split results could be due to the fact that the test participants might have thought that they were expected to exclude the metaphorical items because it was fairly obvious that they do not exist, and were thus just trying to do well in the task.

and was paraphrased as 'to create a time saving shortcut'. These findings will be taken up for discussion in the next section.

4.5 Discussion

The results of the questionnaire study allow for the following conclusions: firstly, novel verbal pseudo-compounds are much more likely to be considered acceptable than genuine verbal compounds. Secondly, the presence of word-families enhances this likelihood even more, as does, thirdly, the presence of a characteristic encoding an unusual procedure that deviates from the norm.

The first criterion dominates the remaining ones (word-families and nameworthiness) since the differences between genuine and derived verbal compounds are the most distinctive. Completely new formations which cannot be related to an established base lexeme appear meaningless and do not lend themselves to plausible interpretation. Only three items among the genuine verbal compounds boast positive scores in the acceptability rating: *to *househop* (arithmetic mean: 0.34 on the scale from +2 to −2; median: 1), *to *spongeclean* (mean: 0.57; median: 1) and *to *fingercomb* (mean: 0.79; median: 2). However, as the results of the questionnaire study reveal, these lexemes benefit from the presence of word-families (cf. e.g. *to barhop*, *to vacuumclean*, *to fingerdry*) and nameworthiness. As will be demonstrated in the following paragraph, such lexemes are formally and semantically dependent on related items and are processed in a way other than compounding.

The presence of word-families, as has been shown, also has a significant impact on the success of a novel verbal compound. Such word-families formally and semantically support novel verbs linked to this network. This support facilitates their understanding and is likely to increase their chances of being accepted, and thus, arguably, their potential for subsequent entrenchment and diffusion (Schmid 2008). The novel verb is associated with already established lexemes and can be interpreted on the basis of something that is already familiar, which is an important way of assigning meaning.

As regards the third aspect, the results demonstrate that novel verbs which denote an activity that deviates from the normal procedure are more likely to be accepted than verbs for concepts which are not particularly nameworthy in this respect. While metaphors also play an important role, a finer differentiation is necessary here. For genuine verbal compounds, which, as the results have shown, tend to defy the assignment of a plausible interpretation anyway, an additional metaphorical element seems to conceal the meaning even further. Since these lexemes, such as *to *weed-sow* (mean: −1.47; median: −2) or *to*

**trust-gamble* (mean: –1.68; median: –2), produced low scores in the acceptability test (overall mean: –1.63; overall median: –2), their chances of being actively used in language are comparatively low. What is interesting, however, is that there are comments in the comprehension part of the study to the effect that, assuming that the word did exist, it would probably have a figurative meaning. Moreover, the paraphrases provided by the participants show parallels to established verbs. This indicates that although the presence of metaphorical elements in completely novel formations negatively influences meaning assignment, they begin to play a crucial role as soon as an anchor (like word-families or a derivation base) is available.

In the following chapter, the results of the tests and their interpretation will be used to develop a model of how verbal compounds are processed by speakers of English, which, in turn, will explain why they do not seem to form a productive schema for verbal compounding in spite of the apparent evidence available to them in the English language.

5 Why Speakers of English do not have a Productive Schema for Verbal Compounding: a Model of How Verbal Compounds are Processed

If speakers were to be able to develop a productive schema for verbal compounding, they would have to process input in such a way that it can serve as evidence for the schema. This would mean that they would have to be able to connect the linguistic forms of verbs apparently formed in a similar way with a set of meanings typically and frequently brought about by this type of formal constellation (Langacker 1987: 492, 2008: 17, Tomasello 2003: 173–175). In short, in order to build a productive schema of verbal compounding, the items constituting the potential input must be processed as compounds. The results of the study indicate, however, that in the processing of verbal compounds language users proceed in a manner that is not conducive to the formation of a verbal compounding schema. When asked to suggest a plausible meaning for a novel verb, participants were found to attempt to revert mentally to established base lexemes or analogous formations which are related via word-families and thus sound similar. A successful meaning assignment, therefore, largely presupposes a base concept which triggers the meaning of the novel verb. Participants do not seem to be able to identify or search for generalized patterns underlying

novel formation, but rely largely on item-specific formal and semantic connections to similar items in the network.

In what follows, we will present several variants of a model which tries to visualize the processes that take place when language users are confronted with different types of novel compound verbs. Examples from the questionnaire study will be discussed below to illustrate the different variants of the "master" model, which will be presented last. In all figures, the stimulus verb is depicted as the rightmost element. Ideas possibly activated in the minds of test participants, which can be gleaned from our test data, are rendered in clouds in some of the later figures.

In the most straightforward case (cf. figure 2), the novel verb, here *to *speed-date*, triggers access to an established base lexeme, here the noun *speed dating*, which evokes the related concept. From this base concept, which denotes 'a process by which people seeking romantic relationships attend organized events at which they have a short conversation with each of several potential partners' (*OED* s.v. *speed dating*, n.), the resulting activity can be derived. Given that this activity is judged as nameworthy – which is assumed to be given here since being pressed for time while dating is rather unusual –,

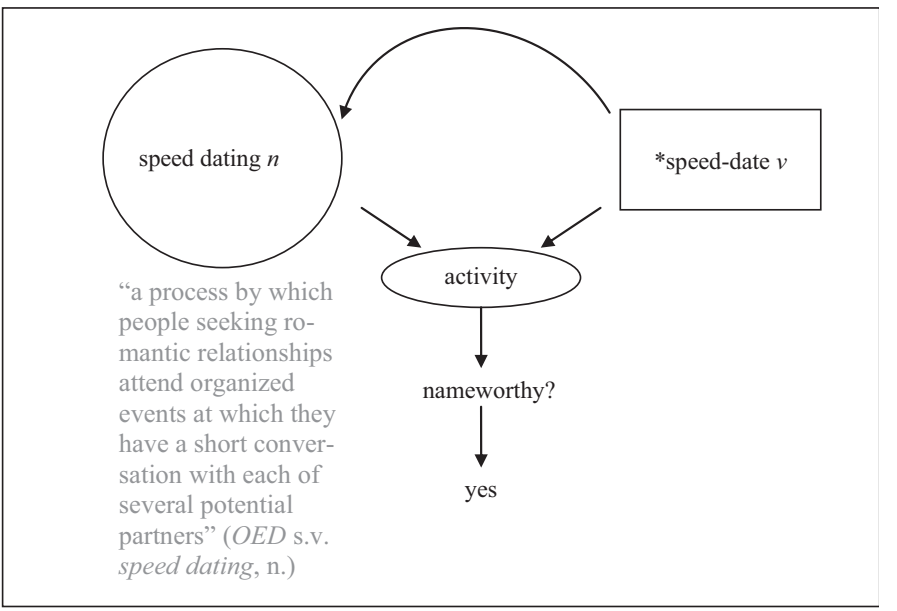

Fig. 2: Model of processing of nameworthy novel verbal compound based on established base lexeme

language users will be satisfied with the result of their processing efforts. In the questionnaire study, this verb yielded very positive results in the acceptability rating task and was unequivocally understood in the derived sense.

The second example (cf. figure 3) deviates from the preceding one in that the derived sense is not nameworthy as such: In this case, the novel verb *to *figure-skate* triggers the established nouns *figure-skating* 'the art or practice of skating in figures' or *figure-skater*, from which a verbal meaning can easily be derived. Although the corresponding verbal compound does not seem to be particularly nameworthy, the acceptability rating score for *to *figure-skate* (1.68; median: 2) is only slightly lower than the one for *to *speed-date* (1.74; median: 2). This may well have to do with potential figurative meanings. As already pointed out above, one participant reacted in a particularly telling way to this impression and offered the paraphrase 'to suggest a solution to a problem that looks/sounds good but is ineffective' as a possible meaning for the verb *to*

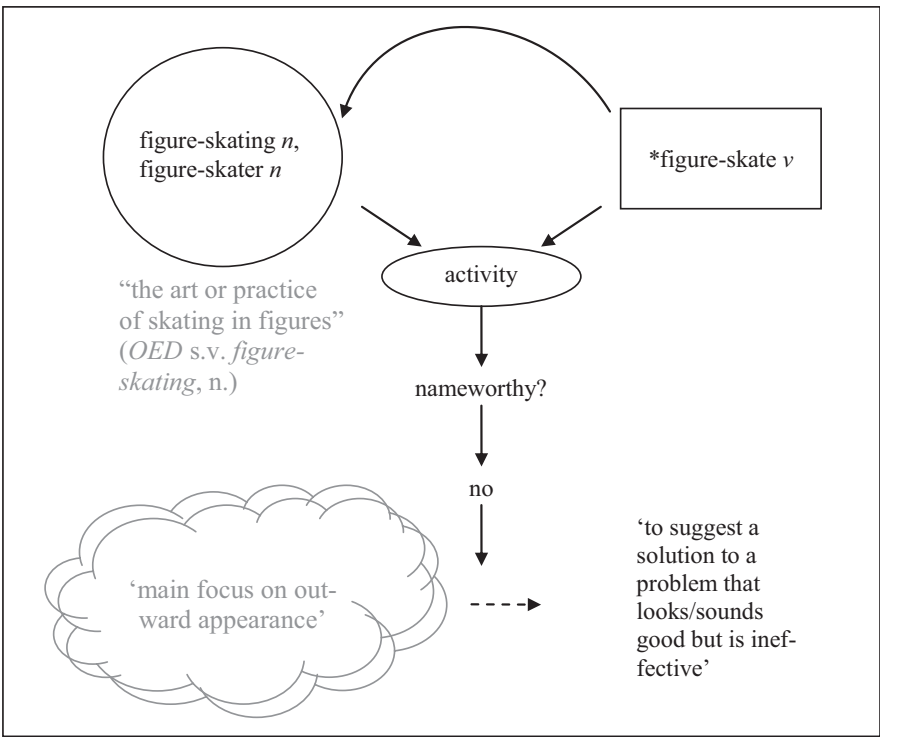

Fig 3: Model of processing of non-nameworthy novel verbal compound based on established base lexeme (partly based on information provided by one test participant)

**figure-skate*. This indicates that the lack of nameworthiness on the literal level can be compensated for by a figurative sense which singles out certain associations related to the nominal concept FIGURE-SKATING and reframes them. The art of figure-skating serves the main purpose of being artistically and visually appealing and thus has a strong focus on the outward appearance, further emphasized by bright dresses and costumes. Mapping these associations onto the domain of troubleshooting results in the hypostatization of a potentially relevant and nameworthy concept. This procedure of creating a figurative meaning next to, or based on, a parallel literal one corresponds to a phenomenon that can generally be observed for verbal pseudo-compounds. In many cases, verbs have both a figurative and a literal meaning which exist side by side. The verb *to sugarcoat*, for example, can be used literally to mean 'to coat with sugar' as well as metaphorically as 'to make something superficially pleasant'.

In the two preceding examples, mental recourse to an established derivation base is possible; the difference concerned the presence or absence of nameworthiness. Such cases of verbal pseudo-compounds stand a good chance of allowing a plausible interpretation based on the underlying nominal concept. A more complex procedure applies to novel genuine verbal compounds where no such derivational base lexeme is accessible and the language user is forced to find another way of assigning a plausible meaning to the verb. In the following example (cf. figure 4), *to *timecut*, it can be seen that related members of word-families can take over the function of providing a conceptual anchor in such cases.

The model in figure 4 illustrates that in cases where no base lexeme is available (**timecutter*/**timecutting* for the novel verb *to *timecut*) the compound verb is split into its constituent elements, here *time* and *cut*. Paraphrases proposed by the test participants such as 'to create a time-saving shortcut' or '[to] reduce the time of something such as a process for efficiency' indicate that the concept SHORTCUT seems to be activated and present while meaning is constructed. This suggests that language users only partly, if at all, try to derive the meaning by computing a semantic relation between the elements *time* and *cut*, which would be typical of nominal and adjectival compounds following well-entrenched formation patterns. Instead, they exploit associations triggered by one of the constituents (here *cut*) to related complex lexemes (here *shortcut*) in deriving a potential meaning. In the course of this, paradigmatic associations to related complex lexemes based on analogy overrule the processing of potential syntagmatic associations between the constituents (cf. Schmid forthcoming). Semantic features like 'reducing time and effort' are thus evoked and transferred to the novel verb *to *timecut*, which explains the associations present in the participants' paraphrases.

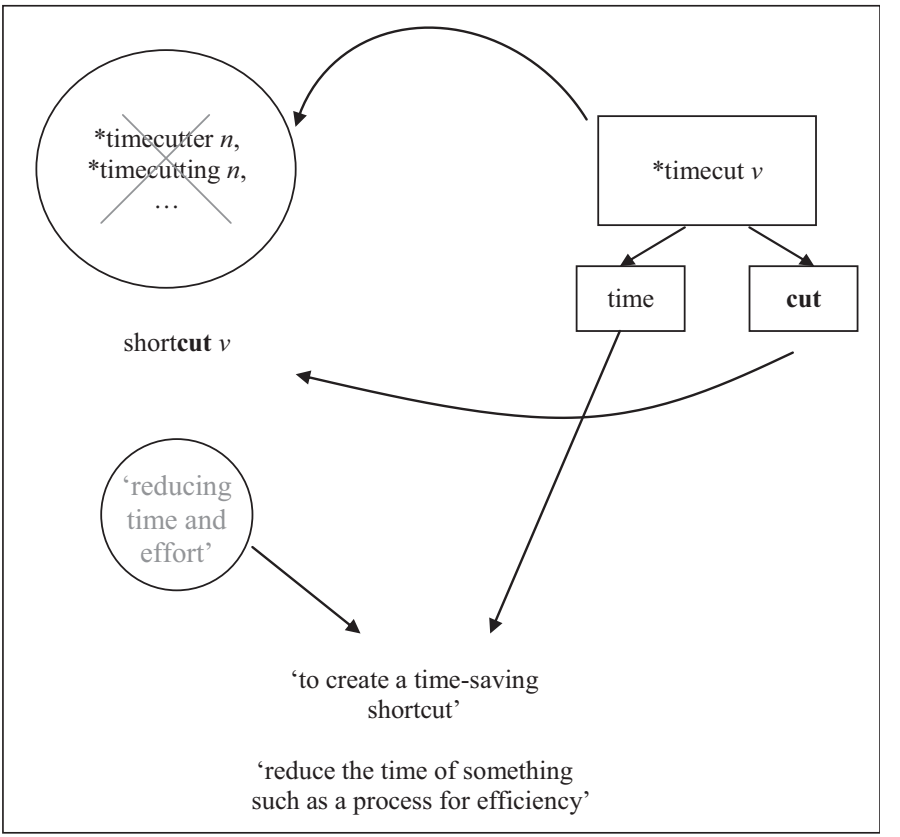

Fig. 4: Model of processing of novel genuine verbal compound (i.e. a compound not based on an established base lexeme) relying on a word-family

The next example (cf. figure 5) resembles the previous one in that no underlying nonverbal base lexeme is available. However, the second constituent *hop* links the stimulus verb *to *household* to already established members of the same word-family of *-hop*-verbs and thus evokes concepts like TABLEHOP, BARHOP and so on. These verbs share the metaphorical meaning 'to move from one place to another', so the paraphrase for the test lexeme offered in the comprehension task reads 'to move from one house to another'. Although the verb yielded fairly positive results, a strikingly high number of participants additionally offered paraphrases which they regarded as more nameworthy and which included associations like having fun, meeting friends, parties, etc., as in 'to move from one house to another, playing and snacking. Either with a group of friends, or to visit friends' or also 'to hop from a party in one house to one in another'. This

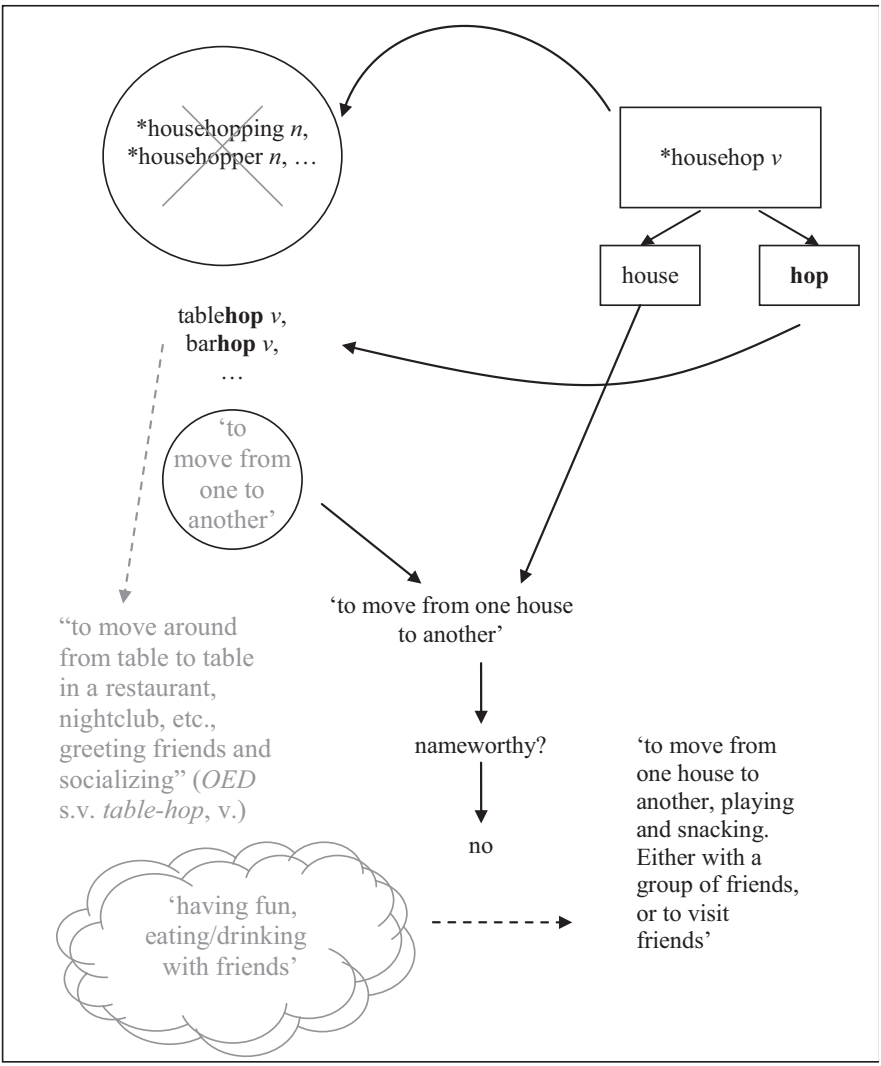

Fig. 5: Model of processing of novel genuine verbal compound (i.e. not based on established base lexeme) relying on a word-family suggesting a metaphorical meaning

is a plausible finding if we consider that established verbs of the same word-family also belong to the category of leisure time. Thus, this example clearly illustrates that word-families can have a significant impact on the meaning of novel verbal compounds in that elements of related concepts, on both a literal and a metaphorical level, can be integrated to yield a nameworthy matter of

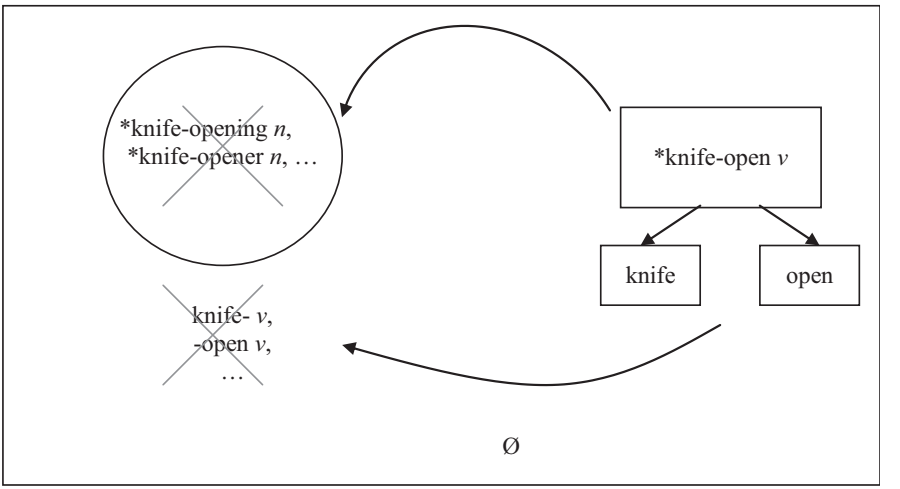

Fig. 6: Model of processing of novel verbal compound not based on established base lexemes and not relying on a word-family

fact. It can be argued that semantically coherent parts of such networks form lower-level schemas (cf. e.g. Langacker 2000: 30–31) which can become increasingly productive as more and more instantiations find a place in the network.

The final example (cf. figure 6) discussed here is concerned with a novel verbal compound (*to *knife-open*) that neither allows recourse to an established base lexeme (**knife-opening/*knife-opener*), nor can it be related to analogous formations via word-families. Since it is also difficult to imagine a plausible nameworthy activity that could be denoted by this verb – in fact none of the participants in the comprehension task was willing to offer a paraphrase – this lexeme does not seem to justifiably exist alongside a syntactic combination and was, therefore, rejected as a possible verbal compound.

The mental processes that are triggered by a confrontation with a novel compound verb are summarized in figure 7 below, which provides a generalized cognitive model of the processing of verbal compounds and pseudo-compounds: when a language user faces a new verbal compound, he or she will, in a first step, attempt to retrieve the meaning by mentally reverting to a typically nominal, more rarely adjectival base concept (1). This means that the verb reminds the language user of related lexemes and triggers corresponding associations. If this attempt fails, the compound will, in a second step (2) be split into its constituents C1 and C2, which might or might not relate the verb to analogous formations via word-family effects. The meanings of items in the network activated by word-family links will be exploited in the construction of a potential meaning of the target, especially if the envisaged concept appears to be

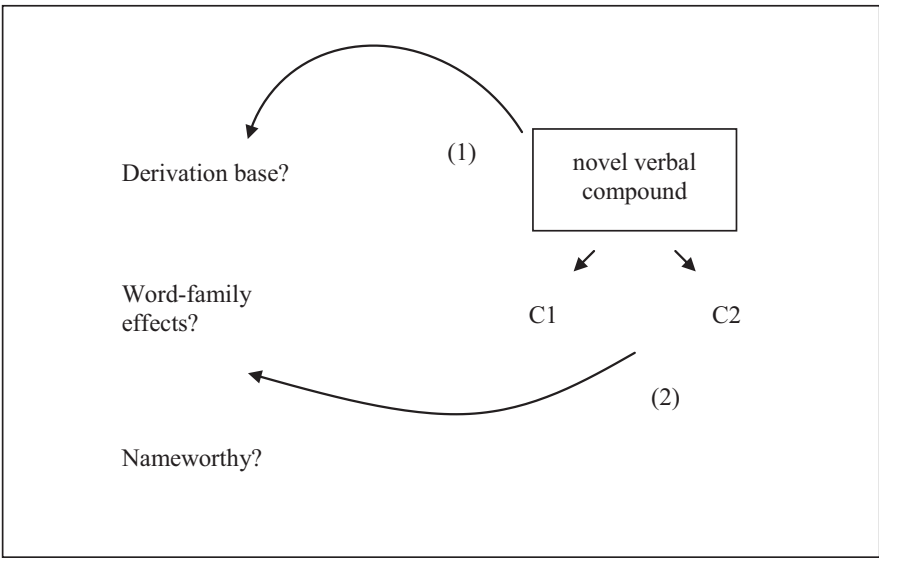

Fig. 7: Generalized model of processing of novel genuine verbal compounds and novel pseudo-compounds

nameworthy to some degree. The more easily a plausible meaning can be derived, the more likely the novel compound will be accepted by the speaker.

6 Conclusion

The key question introduced at the beginning of this paper concerned the nature of the general knowledge that speakers of English are likely to have about verbal compounding, which presumably influences the chances of whether or not individual verbal compounds are entrenched in the minds of language users and established in the lexicon. Based on an initial analysis of established verbal pseudo-compounds from two different sources, the resulting hypotheses concerning the acceptability and comprehensibility of compound verbs were tested by means of a detailed questionnaire study. It has been shown that some fabricated compound verbs like *to *househop* or *to *fingercomb* yield rather positive results indicating that native speakers could imagine them being used. A closer inspection based on the results of the empirical analyses revealed, however, that these lexemes are not processed cognitively by applying a compounding schema, but are rather split into their constituents and matched against familiar

lexemes and thus interpreted by means of analogy to established verbs of the same word-family. In this process, the question of whether or not a given verbal compound denotes a nameworthy concept due to some kind of norm deviation or a metaphorical element plays a key role.

This suggests that speakers of English do not coin genuine verbal compounds, even though the English language includes a number of lexemes which superficially resemble them, because as hearers they actually do not decompose and combine these forms in the way they would process typical nominal or adjectival novel compounds. The meaning paraphrases for novel compound verbs that participants provided as part of the questionnaire study indicate that existing base lexemes and word-families strongly influence the acceptability of new formations. Strictly speaking, however, complex verbs whose meanings are assigned on the basis of an analogy to related words only should not be considered as genuine compounds, precisely because a productive schema does not seem to be involved. As illustrated with the examples *to *household* and *to *timecut* above, these verbs cannot be said to be processed independently of their word-families because related lexemes are required to be active in the speaker's mind at the moment they receive their interpretation.

As illustrated by the model discussed in the preceding chapter, it can be claimed that there is a decently routinized procedure for processing novel verbal pseudo-compounds, which – provided that a nameworthy matter of fact is denoted – can be derived readily from their nominal bases. Whether this can be considered a productive schema is doubtful, however. What is clear is that such a schema is not available for genuine verbal compounds. Whereas novel verbs which can rely on the positive effects of word-families were found to be acceptable at least to a certain degree, genuine verbal compounds, which do not evoke any related lexemes, were perceived as extremely peculiar and not qualified to receive a plausible interpretation that could justify their existence.

Thus, to date, verbal compounding remains a non-productive process in English. If it can be claimed to be productive to some extent, then this highly limited productivity does not seem to arise from the existence of a rule or general schema, but is based on links in the network and potential lower-level schemas emerging from them, on the analogy to existing formations as well as, most importantly, the support of existing nouns from which verbal pseudo-compounds can be derived. The evidence collected here thus salvages the basic assumptions of the usage-based approach in that it explains why a schema is not formed in spite of the evidence seemingly available. In addition, our findings indicate that the network idea should play a key part in models of knowledge of at least word-formation, possibly also lexicon and syntax, as the tendency of

language users to take recourse to other elements in the network and to exploit their semantic properties is one of the key insights of this study.

Works Cited

Ackema, Peter and Ad Neeleman. 2004. *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.

Adams, Valerie. 1973. *An Introduction to Modern English Word-Formation*. London: Longman.

Adams, Valerie. 2001. *Complex Words in English*. Harlow: Pearson Education Limited.

Barlow, Michael and Suzanne Kemmer (eds.). 2000. *Usage-Based Models of Language*. Stanford, CA: CSLI Publications.

Booij, Geert. 2005. *The Grammar of Words*. Oxford: Oxford University Press.

Brömser, Bernd. 1985. "On the Derivation of English Verbal Compounds". In: Wilfried Kürschner and Rüdiger Vogt (eds.). *Grammatik, Semantik, Textlinguistik: Akten des 19. Linguistischen Kolloquiums Vechta 1984. Band 1.* Tübingen: Niemeyer. 99–113.

Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge/New York: Cambridge University Press.

Cho, See-Young. 2002. "Synchrone und diachrone Untersuchungen zu den zusammengesetzten Verben im Englischen". Habilitationsschrift, Technische Universität Berlin.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.

COCA = The Corpus of Contemporary American English: 450 Million Words, 1990–Present. 2008–. Ed. Mark Davies. <http://corpus.byu.edu/coca/>.

Downing, Pamela. 1977. "On the Creation and Use of English Compound Nouns". *Language* 53: 810–842.

Fillmore, Charles J. 1968. "The Case for Case". In: Emmon Bach and Robert T. Harms (eds.). *Universals in Linguistic Theory.* New York etc.: Holt, Rinehart and Winston. 1–88.

Grimm, Jacob. 1877. *Deutsche Grammatik: Zweiter Teil*. Neuer vermehrter Abdruck. Gütersloh: Bertelsmann.

Hall, Robert A. 1956. "How We Noun-Incorporate in English". *American Speech* 31: 83–88.

Hansen, Barbara, Klaus Hansen, Albrecht Neubert und Manfred Schentke. 1990. *Englische Lexikologie: Einführung in Wortbildung und Lexikalische Semantik*. 2nd ed. Leipzig: Enzyklopädie.

de Jong, Nivja H. M., Laurie B. Feldman, Robert Schreuder, Matthew Pastizzo and R. Harald Baayen. 2002. "The Processing and Representation of Dutch and English Compounds: Peripheral Morphological and Central Orthographic Effects". *Brain and Language* 81: 555–567.

Kastovsky, Dieter. 1986. "Diachronic Word-Formation in a Functional Perspective". In: Dieter Kastovsky and Aleksander J. Szwedek (eds.). *Linguistics across Historical and Geographical Boundaries. In Honour of Jacek Fisiak on the Occasion of his 50th Birthday.* Volume 1: *Linguistic Theory and Historical Linguistics*. Berlin: Walter de Gruyter. 409–422.

Kerremans, Daphné. 2012. "A Web of Words: On the Conventionalization of English Neologisms". Dissertation, Ludwig-Maximilians-Universität München.

Kirchner, Gustav. 1959. "'Neue Synthese' im Gegenwartsenglisch". In: Horst Oppel (ed.). *Festschrift für Walther Fischer*. Heidelberg: Winter. 302–322.

Lamberty, Angela. 2012. "Why we Don't *Cardrive* or *Bookread*, but *Slavedrive* and *Lipread:* A Cognitive-Linguistic Approach to Verbal Compounds and Pseudo-Compounds in English". Dissertation, Ludwig-Maximilians-Universität München.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. Volume 1: *Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 1988. "A Usage-Based Model". In: Brygida Rudzka-Ostyn (ed.). *Topics in Cognitive Linguistics*. Amsterdam/Philadelphia: Benjamins. 127–161.

Langacker, Ronald W. 2000. "A Dynamic Usage-Based Model". In: Michael Barlow and Suzanne Kemmer (eds.). *Usage-Based Models of Language*. Stanford, CA: CSLI Publications. 1–63.

Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford/New York: Oxford University Press.

Leech, Geoffrey N. 1981. *Semantics*. 2nd ed. Harmondsworth: Penguin.

Leisi, Ernst. 1975. *Der Wortinhalt: Seine Struktur im Deutschen und Englischen*. 5th ed. Heidelberg: Quelle & Meyer.

Lieber, Rochelle. 2009. "IE, Germanic: English". In: Rochelle Lieber and Pavol Štekauer (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 357–369.

Lipka, Leonhard. 1977. "Lexikalisierung, Idiomatisierung und Hypostasierung als Probleme einer synchronischen Wortbildungslehre". In: Herbert E. Brekle and Dieter Kastovsky (eds.). *Perspektiven der Wortbildungsforschung: Beiträge zum Wuppertaler Wortbildungskolloquium vom 9.–10. Juli 1976. Anläßlich des 70. Geburtstags von Hans Marchand am 1. Oktober 1977*. Bonn: Bouvier. 155–164.

Lipka, Leonhard. 1983. "A Multi-Level Approach to Word-Formation: Complex Lexemes and Word Semantics". In: Shirô Hattori and Kazuko Inoue (eds.). *Proceedings of the 13th International Congress of Linguists: August 29 – September 4, 1982, Tokyo*. Tokyo: The Committee. 926–928.

LDOCE = Longman Dictionary of Contemporary English. 2005. 4th ed. by Della Summers. Harlow: Pearson Education. (CD-ROM).

Marchand, Hans. 1969. *The Categories and Types of Present-Day English Word-Formation. A Synchronic–Diachronic Approach*. 2nd ed. München: Beck.

Mithun, Marianne. 1984. "The Evolution of Noun Incorporation". *Language* 60: 847–894.

Mithun, Marianne. 1985. "Diachronic Morphologization: The Circumstances Surrounding the Birth, Growth and Decline of Noun Incorporation". In: Jacek Fisiak (ed.). *Papers from the VIth International Conference on Historical Linguistics, Poznán, 22–26 August 1983*. Current Issues in Linguistic Theory 34. Amsterdam/Philadelphia: Benjamins. 365–394.

Mithun, Marianne. 1986. "On the Nature of Noun Incorporation". *Language* 62: 32–37.

OED = Oxford English Dictionary. 2000–. 3rd ed. online. Oxford: Oxford University Press. <http://www.oed.com/>.

Pennanen, Esko V. 1966. *Contributions to the Study of Back-Formation in English*. Tampere: Yhteiskunnallinen Korkeakoulu.

Schmid, Hans-Jörg. 2008. "New Words in the Mind: Concept-Formation and Entrenchment of Neologisms". *Anglia* 126: 1–36.

Schmid, Hans-Jörg. 2011. *English Morphology and Word-Formation: An Introduction*. Berlin: Schmidt.

Schmid, Hans-Jörg. forthcoming. "Lexico-Grammatical Patterns, Pragmatic Associations and Discourse Frequency". In: Thomas Herbst, Hans-Jörg Schmid and Susen Faulhaber (eds.). *Constructions – Collocations – Patterns*. Berlin etc.: Mouton de Gruyter.

Schreuder, Robert and R. Harald Baayen. 1997. "How Complex Simplex Words Can Be". *Journal of Memory and Language* 37: 118–139.

Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA/London: Harvard University Press.

de Vaan, Laura, Robert Schreuder and R. Harald Baayen. 2007. "Regular Morphologically Complex Neologisms Leave Detectable Traces in the Mental Lexicon". *The Mental Lexicon* 2: 1–23.

Vendler, Zeno. 1957. "Verbs and Times". *The Philosophical Review* 66: 143–160.

Zimmer, Karl E. 1971/1981. "Some General Observations about Nominal Compounding". In: Leonhard Lipka and Hartmut Günther (eds.). *Wortbildung*. Darmstadt: Wissenschaftliche Buchgesellschaft. 233–257. [Reprint of "Some General Observations about Nominal Compounds". *Working Papers on Language Universals* 5. Stanford University, CA: Committee on Linguistics].

Appendix

Test stimuli for acceptability rating and comprehension task and overall results of acceptability rating

Genuine compounds	Word-family effects	Name-worthiness	Median of acceptability rating	Arithmetic mean of acceptability rating
airstroll	yes	metaphor	−2	−1.80
airtest	yes	no	−1	−0.26
bellykick	yes	norm deviation	−1	−0.20
cardpay	no	no	−2	−0.80
charm-snare	no	metaphor	−2	−1.68
coldeat	yes	no	−2	−1.81
coldvisit	yes	metaphor	−2	−1.14
colourcook	yes	norm deviation	−2	−1.49
colourtaste	yes	no	−2	−1.50
couchsleep	no	norm deviation	−1	−0.71
crutchwalk	yes	norm deviation	−2	−1.50
eyeread	yes	no	−2	−1.39
fear-bleed	no	metaphor	−2	−1.74
fingercomb	yes	norm deviation	2	0.79
floorsit	yes	norm deviation	−1	−0.95
flypick	yes	no	−2	−1.06
friendpile	no	metaphor	−2	−1.59
hand-eat	yes	no	−1	−0.68
headpeck	yes	metaphor	−2	−1.17
headplunge	yes	norm deviation	−1	−0.50
househop	yes	metaphor	1	0.34
knife-open	no	no	−1	−0.94
mashfeed	yes	no	−2	−1.49
massarrive	yes	no	−2	−1.44
rumour-spread	no	no	−1	−0.38
schoolhop	yes	metaphor	−1	−0.30
shame-lie	no	no	−2	−1.28
spongeclean	yes	no	1	0.57
stick-discipline	no	norm deviation	−2	−1.42
table-eat	no	no	−2	−1.54
timecut	yes	metaphor	−2	−1.04
trust-gamble	no	metaphor	−2	−1.68
watertest	yes	no	−1	−0.04
weed-sow	no	metaphor	−2	−1.47
windowcheck	yes	no	−2	−1.04

Pseudo-compounds	Word-family effects	Name-worthiness	Median of acceptability rating	Arithmetic mean of acceptability rating
air-freshen	yes	no	1	0.90
beauty-sleep	no	norm deviation	1	1.14
clod-hop	yes	metaphor	1	0.11
comfort-eat	no	norm deviation	1	0.28
curtain-raise	no	metaphor	–1	–0.35
earshoot	yes	metaphor	–2	–1.70
fabric-soften	no	no	2	1.06
face-save	yes	metaphor	–1	–0.44
figure-skate	yes	no	2	1.68
fire-eat	yes	metaphor	1	0.02
food-poison	no	norm deviation	1	0.38
foot-drag	no	metaphor	–1	–0.18
garden-party	no	no	–1	–0.25
guestwork	yes	no	–2	–1.19
hand-kiss	yes	no	1	0.15
hand-signal	yes	no	2	1.42
hand-stamp	yes	no	2	1.33
handstand	yes	norm deviation	2	1.14
homespin	yes	metaphor	–1	–0.59
laser point	yes	no	1	0.22
lion-tame	no	metaphor	2	1.07
mudbathe	yes	norm deviation	2	1.30
nametape	yes	no	–2	–1.03
palm-read	yes	metaphor	2	1.75
pillsleep	no	norm deviation	–2	–1.68
potato peel	no	no	–1	–0.72
purpose-build	no	no	–1	–0.10
question-fire	yes	metaphor	–2	–1.02
shotgun-marry	no	metaphor	1	–0.14
side order	yes	no	1	0.51
speed-date	yes	norm deviation	2	1.74
stamp-collect	no	no	2	0.97
sticker-price	no	no	–1	–0.92
stickwalk	yes	norm deviation	–1	–0.84
stone-wash	yes	norm deviation	2	1.44
window-clean	yes	no	2	0.94

Distractors	Median of acceptability rating	Arithmetic mean of acceptability rating
cherrypick	2	1.74
handwash	2	1.79
sunbathe	2	1.76