NEW WORDS IN THE MIND: CONCEPT-FORMATION AND ENTRENCHMENT OF NEOLOGISMS

Abstract: So far, the study of new words and the early stages of their lexicalization and institutionalization has focused very much on the structural and semantic changes involved as well as on the gradual spread of words in a speech community. This paper focuses on insights into the concomitant processes taking place in language users' minds. It takes up ideas on concept-formation and hypostatization put forward in the philosophy of language, word-formation and lexical semantics and relates them to recent evidence on the processing and storage of nonce-formations and recently coined complex words collected by psycholinguists and neurolinguists. The role of frequency of exposure and semantic transparency in the increasing entrenchment of concepts in language users' memories is discussed. Effects of hypostatization – the subjective impression that the existence of a word suggests the existence of a class of things denoted by the word – are shown to be very strong even in the early phases of the establishment of new words, and pragmatic exploitations of these effects are explored.[1]

1. Introduction

This paper is concerned with new words that are made up from existing morphological material (rather than so-called *creations ex nihilo*). It discusses cognitive aspects of the early phases in the establishment of *neologisms*.[2] This period begins with the first use as nonce-formations, its end is

[1] I would like to thank Harald Baayen, Maura Bresnan-Enders, Wolfgang Falkner, Peter Hohenhaus, Lucia Kornexl, Annette Mantlik and Hans Sauer for their invaluable comments on earlier versions of this paper.

[2] Neologisms are not simply 'new words'. Rather, at least in theoretical terms, they are words which have lost their status as nonce-formations and are in the process of becoming or already have become part of the norm of the language (Bauer 2001, 39; Lipka, Handl & Falkner 2004, 3), but are still considered new by most members of a speech community (Fischer 1998, 3; Hohenhaus 2005, 364). They are thus inherently transitional phenomena (Hohenhaus 2005, 365). This of course implies that a word may be a neologism for one language user and familiar to another, and that in the absence of clues provided by the speaker signalling the newness of the word (see Section 6.1.1. below), hearers will be unsure whether they are confronted with a new word or an existing word unfamiliar to them (cf. Hohenhaus 2006, 17).

DOI 10.1515/angl.2008.002

marked by various symptoms of an increasing and successful entrenchment in the minds of speakers of the language and concomitant institutionalization in the speech community (or considerable sections thereof). It is common in linguistics to assume that this development is accompanied by a growing degree of lexicalization, i.e. the emergence of specific semantic, orthographic, phonological or syntactic properties of the complex lexeme, which require it to be listed as an entry in its own right in the lexicon of the language.

The structural and sociolinguistic changes involved in this development have been described and explained in considerable detail in a number of publications including Bauer (1983, 42–61 and 2001, 33–47), Kastovsky (1982, 164–168), Lipka (1992), Sauer (1992), Hohenhaus (1996 and 2005), Fischer (1998) as well as Lipka, Handl & Falkner (2004). The focus of this paper will be on the processes taking place in the minds of language users processing nonce-formations and gradually acquiring neologisms.

2. Background: Three Perspectives on and Three Stages of Establishment

In spite of, or perhaps rather as a result of the undeniable interest taken in this topic by many specialists, conceptual and terminological confusion has dominated the area under consideration here, as it has so many other well-researched fields in linguistics. In yet another attempt to structure the multiplicity of processes that can be distinguished in the development of new words towards establishment and in order to increase terminological transparency,[3] I have introduced a distinction between three perspectives on the development, on the one hand, and three stages of it, on the other (Schmid 2005, 71–85):

1. the *structural* (perspective on the development of the properties of the word itself;
2. the *socio-pragmatic* perspective on the spread of familiarity of a word in a speech community;
3. the *cognitive* perspective on the formation and entrenchment of a concept associated with the word in the minds of the members of a speech community.

The processes brought into view by the three perspectives are subsumed under the labels

[3] For recent state-of-the-art accounts of the terminology in the area see Bauer (2001, 43–47), Lipka, Handl & Falkner (2004), Hohenhaus (2005).

1. *Lexicalization* (structural perspective),
2. *Institutionalization* (socio-pragmatic perspective), and
3. *Concept-formation* (cognitive perspective) respectively.

The superordinate term for the whole development is *establishment* (cf. Bauer 2001, 46). The continuum from the first use to complete establishment is, admittedly somewhat arbitrarily, carved up into three stages referred to as *creation, consolidation* and *establishing*. Each of the three perspectives highlights different aspects of these three phases, which are termed as follows:

1. Structural perspective: (product of) nonce-formation, stabilization and lexicalized lexeme.
2. Socio-pragmatic perspective: (process of) nonce-formation, spreading and institutionalized lexeme.
3. Cognitive perspective: pseudo-concept, process of hypostatization and hypostatized concept.

The conceptual and terminological system is summarized in Table 1 below. Unfortunately, the table suggests that the processes arranged in one horizontal line of cells always take place at the same time. This is not the case, however. For example, it is not unlikely that the process of hypostatization is temporally prior to and logically necessary for the spreading of a new word.

Table 1: Three perspectives and three stages of the establishment of new words

Perspectives: / Stages:	Structural perspective	Socio-pragmatic perspective	Cognitive perspective
creation	(product of) nonce-formation	(process of) nonce-formation	pseudo-concept
consolidation	stabilization	spreading	(process of) hypostatization
establishing	lexicalized lexeme	institutionalized lexeme	hypostatized concept

As already mentioned in Section 1 above, the focus of this paper is on the column on the right-hand side of the table. Concerning the structural and socio-pragmatic perspectives, I will restrict my account to the following brief summary.[4]

[4] References to important sources can be found in Section 1 above.

3. Lexicalization and Institutionalization

With regard to both formal and semantic aspects, the process of lexicalization has been described as a development beginning with the creation of a nonce-form which serves to express a new meaning. In this birth situation, context and co-text tend to reduce the ambiguity of the intended meaning to such an extent that it can generally be considered more or less clear-cut for both the initial speaker and his or her hearers. For language users confronted with the new word in different contexts later during its spread, this may no longer be the case. As a result of this indeterminacy, both the meaning of the neologism, which at this stage is still context-dependent and ambiguous, and its orthographic and phonological form may be subject to modifications by new users. As the word gradually becomes institutionalized, i. e. gains wider acceptance and becomes item-familiar to more and more speakers, form and meaning stabilize. Ambiguity and context-dependence are reduced and the lexeme tends to develop semantic autonomy and context-independence, so that speakers can effortlessly recognize and identify its meaning. The semantic aspects of completed lexicalization are captured by notions such as *idiomaticity*, *opacity*, *lack of transparency* or *compositionality*, all referring to the fact that the meaning of the composite form is no longer computable on the basis of the meanings of the constituents and the semantic relation typically expressed by the word-formation pattern applied.[5]

One important point must be emphasized here. Lexicalization is usually considered a diachronic process whose effects can also be recognized in synchronic sections. However, diachronic change may well be less important for semantic lexicalization than is generally assumed when it is exemplified with stock examples of opaque compounds (cf. Faiß 1978) such as *lord* (< OE *hlāf-weard* 'loaf keeper'), *gospel* (< *gōd-spel* 'good news') or, still transparent but nevertheless semantically idiosyncratic, *holiday* (< *hālig-dæg* 'holy day'). Lipka (1981, 122) has drawn attention to the fact that many complex lexemes are non-compositional from the very start and has termed formations of this type *instantaneous coinings*. Handl (1999) carried out a large-scale dictionary study tracing the lexicalization and institutionalization development of 396 neologisms coined between 1958 and 1973. The words were recruited from two dictionaries of neologisms (Green 1991 and Barnhart, Steinmetz & Barnhart 1980). Their semantic development and degree of institutionalization was monitored with the help of

[5] Note that the notions of *lack of transparency* and *lack of compositionality* are not synonymous. Complex lexemes can easily be transparent without being entirely compositional. For example, a simple compound like *barman* is undoubtedly transparent even without a context, but the paraphrase 'a man working in a bar' shows that it is not entirely compositional because of the element 'working', not to mention additional aspects of its lexicalized conceptual content, such as 'mixes drinks' and 'serves drinks'.

dictionaries from the Oxford family, viz. the sixth and the ninth edition of *The Concise Oxford English Dictionary of Current English* (Sykes 1976 and Thompson 1995) as well as *The New Oxford Dictionary of English* (Pearsall 1998). Handl (1999, 98) found that no less than 39 % of the neologisms had idiomatic meanings at their very birth. And even the remaining words included considerable extra semantic content in addition to their strictly compositional meanings. For example, the meaning of the compound *noise pollution*, which was rated non-idiomatic by Handl, undoubtedly goes beyond the meanings of the constituent nouns: 'production of noise of motor vehicles, jet planes etc, viewed as harmful to people and the environment' (Handl 1999, 88).[6]

4. Conceptual Aspects: Two Approaches

Cognitive aspects of the establishment of new words in speakers' minds have essentially been investigated in two types of frameworks working more or less independently of each other (cf. Schmid 2005, 73–85). The first (dealt with in Section 5) has its roots in language philosophy and linguistic semantics. Essential to this approach is the notion of *hypostatization*, which captures the impression that the existence of a word suggests the existence of a thing or entity denoted by the word (Lipka 1977, 161). The origins of the second framework (Section 6) can be found in psycholinguistics. This approach relies on the metaphor of the mental lexicon as a module of the mind where knowledge about words is stored in the form of a huge network and accessed and activated when required. A corollary image favoured in cognitive linguistics is the entrenchment of cognitive, and possibly neurological, routines that are experienced as concepts in our subjective consciousness.

5. Hypostatization

Linguists and language philosophers have been concerned with the subjective states of mind experienced by language users when they produce and understand words. Their main focus has been on what Leech (1981, 32) termed the "concept-forming power of the word". This notion accounts for the naïve but deeply ingrained impression that words stand for concepts. An even stronger claim is that the existence of a particular word creates the impression that there is a corresponding thing or entity to which the word refers. This effect is known as *hypostatization.*

---

[6] For a more detailed discussion of this phenomenon from a cognitive point of view see Section 6.1.1. below.

The *locus classicus*, at least in German-speaking English linguistics, is undoubtedly Leisi's account of *hypostatization:*

> Die Mythologie, der scholastische Realismus und die platonische Ideenlehre sind die größten Beispiele für die Neigung der Sprachgemeinschaften, jede Erscheinung irgendwelcher Art, sofern sie durch ein Wort bezeichnet werden kann, zu vergegenständlichen (allenfalls zu personifizieren) und mit einer selbständigen, von anderen Erscheinungen abgelösten Existenz zu begaben, sie also zur akzidenzlosen Substanz zu erheben. Diese Erhebung zur Substanz nennen wir, dem Sprachgebrauch der Philosophie folgend, Hypostasierung. (Leisi 1975, 26; emphasis original)

It is important to note that the phenomenon of hypostatization is not simply a metaphysical speculation but has strong and observable effects on how we perceive the world. Leisi refers to Cassirer's work on mythology and language and to Chase's critical account of how the reification of abstract notions influences our world view.[7] Criticism of the hypostatizing effects of words is also prominent in the so-called *General Semantics* movement of the 1940s and in Bolinger's important book *Language: The Loaded Weapon* (1980).

Hypostatization is achieved by all established content words. What is more important for our discussion here is that the effect can also be observed for neologisms and is thus crucial for the understanding of word-formation. This has been highlighted by Lipka, who defines the notion as follows:

> Unter Hypostasierung durch das Wort verstehe ich die Erscheinung, daß die Existenz eines sprachlichen Zeichens auch die Existenz eines einzigen von diesem bezeichneten Dings suggeriert. Diese Suggestion [...] bewirkt eine Vergegenständlichung, eine Erhebung zur Substanz. (Lipka 1977, 161)

While this rightly emphasizes the reifying effects of words, it should be added that it is the existence of *a whole class of referents* that is evoked rather than *one single individual* (Hohenhaus 1996, 317).

Although the power to hypostatize is inherent in all types of (content) words, the strength and salience of the effects of the process differ. As far as the different word-classes are concerned, nouns have a stronger hypostatization potential than other parts of speech. This is due to the fact that nouns tend to profile conceptual content as 'things',[8] as object-like entities with neat boundaries in space and a stable existence across time. This property of nouns makes them particularly good candidates for reification. Comparing nouns to adjectives, Bolinger (1980, 79) makes this very clear and empha-

---

[7] For more references see Leisi (1985, 253–257 and 268).

[8] See Langacker (1987, 189) and, much earlier, Gardiner (1932, 9f.) and Leisi (1975, 26f.).

sizes the effect that a familiar cognitive category of entities is apparently invoked by the use of a noun:

> The noun OBJECTIFIES in a way the adjective cannot. A quality may come and go. If we are disappointed at Jane's lack of appreciation we can call her *ungrateful*, or solidify it a step further and call her *an ungrateful person*. But if we call her an *ingrate* we put a brand on her: the noun implies that the world puts people like this in a class by themselves.

Even within the word-class of nouns, however, hypostatization does not always play the same role. This can be shown by comparing three types of nouns, i.e. concrete nouns, nouns denoting events, actions and activities, and abstract nouns. Examples used for illustrative purposes are recent neologisms listed in the archives of the Macmillan *Word-of-the-Week* website[9] and in Maxwell (2006): the concrete noun *she-pee*, the event noun *booze-cruise* and the abstract noun *infomania*.

The concrete noun *she-pee* has both a phonological and a semantic motivation: it is formed with a phonological allusion to the rhyming *tepee*, and its constituents point to two crucial semantic aspects related to its meaning:

> a tent containing urinals (open toilets which are usually fixed to a wall), designed specifically for use by women at open-air events. (Macmillan *Word-of-the-Week* website)

The role of hypostatization is hardly noticeable in this recent addition to the English lexicon, since the word denotes a class of bounded objects. It is not inconceivable that the inventors of the object had distinct ideas about the design of these facilities and their purpose before they came up with a suitable name (cf. Lipka 1981, 125). Very much along the lines suggested by Platonic realism, the word apparently serves simply as a label for a pre-existing (idea of) a category of things.

This line of reasoning is not convincing for the other two N+N-compounds, *booze-cruise* denoting an event-concept, and *infomania* (with a clipped first constituent) denoting an abstract concept.[10] To be sure, the phenomena denoted by these two words most likely also existed before the words were coined: trips from the UK to the Continent with the main or sole purpose of buying alcoholic beverages cheaply (*booze-cruise*), and the tendency to be distracted from one's real work by continually responding to electronic communication (*infomania*). What is crucial, however, is that it is

---

[9] URL: <http://www.macmillandictionary.com/resourcenew-archive.htm>.

[10] An alternative analysis of *infomania* would consider *-mania* as a so-called final combining form (cf. Bauer 1983, 213–216; Bauer & Huddleston 2002, 1661f.) also in evidence, e.g., in established neo-classical formations like *kleptomania*, *nymphomania* or *megalomania*. Fischer (1998, 153–159) also considers elements such as *info-* as combining forms and investigates symptoms of their gradual institutionalization in *The Guardian*.

unlikely that speakers of English had the concepts of the phenomena prior to the existence of these terms.[11] This is not to say that speakers could not have been aware of the phenomena as such. But they would not have experienced these phenomena as manifestations of recurrent and familiar events or personal habits, mental representations of which seem to pop up readily in their minds. To the language user, the nouns carve an apparently neatly bounded segment from the constant flux of events going on in the world around them. This is what the impression of having a concept of something is all about.

In the terms familiar from the semiotic triangle, the concept associated with *she-pee* is essentially motivated and sustained by knowledge about the referents. The concepts linked to *booze-cruise* and *infomania*, on the other hand, depend very much on the symbol (i.e. the word) for their inception and support. Here it is the word that turns a rather diffuse network of related ideas – e.g. 'travel', 'the Continent, France', 'buy', 'drink', 'cheap', 'alcohol' – into one holistic and integrated conceptual unit. Similarly, many people may have a feeling that they themselves as well as their colleagues are too busy answering emails and searching for information on the internet to get their real work done. This feeling immediately turns into a gestalt-like concept, a "piece of frozen reality" (Brekle 1978, 75; Kastovsky 1982, 155), once these people learn that there is a word for it. "Configurations of qualities which have previously not been objectified go through a process of objectification and are perhaps raised to the level of a new entity" (Brekle 1978, 75). In sum, whereas the concrete noun seems to do no more than label or name a class of things, the event noun and the abstract noun actually create a concept.

A special case of hypostatization are concrete nouns with fictive referents. These provide a particularly suitable testing-ground for the effects of hypostatization (see Hohenhaus 1996), because they demonstrate how 'things' can be created by the creation of words. Referring to science fiction novels, Kastovsky (1978, 360) makes the following observation:

> Die technische Staffage, die diese Romane zu charakterisieren pflegt – und die es in aller Regel ja gar nicht gibt –, wie z.B. *matter converter, star-drive, hyper-drive, stunner, blaster, space warp* usw., wird praktisch nie deskriptiv eingeführt, sondern es wird, ihre Existenz voraussetzend, durch Nennung der betreffenden Bezeichnung direkt auf sie referiert.

[11] Obviously this presupposes that established words with identical or similar meanings did not exist prior to the coining of *booze-cruise* and *infomania*. According to the online edition of the *OED*, the lexeme *booze-cruise* already existed in the United States during the time of prohibition, with people boarding ships taking them "just far enough from the shoreline to be outside U.S. jurisdiction, so that they could buy and consume alcohol legally" (*OED* online, s.v. *booze-cruise*).

The effects of hypostatization are also exploited in advertising discourse, where new names are coined and hammered into consumers' heads to insinuate the invention of new products or new qualities of familiar products.

Another exploitation of the hypostatization of new words is the promotion of a new business idea or field of expertise. For example, in an endeavour to market her ideas on *container gardening*, the British horticulturalist Marjorie Mason Hogue coined the noun *potscaping* ('the artistic arrangement of flowers and shrubs planted in pots and other containers', Macmillan *Word-of-the-Week* website), presumably on the model of *landscaping*. It is no coincidence that this new practice was couched in terms of a nominalization, since, as we have seen, the noun has a more powerful hypostatization potential. Event-nominalizations using the suffix *-ing*, also highlighted by Bolinger (1980, 63), are in fact conspicuously frequent in the list on the Macmillan *Word-of-the-Week* website, which includes *extreme ironing, shopgrifting, speed dating* and *speed networking, egosurfing, warchalking, stagephoning* and *nanopublishing*.[12] In Bolinger's words, "if an existing name is a certificate of a thing, the making of a new name is a certificate for the making of a new thing" (1980, 63).

What all this suggests, then, is that hypostatization is a general effect of the use of all words, no matter whether they are familiar or new to the hearers/readers. With this insight in mind, I will now turn to psycholinguistic and neurolinguistic evidence to see whether this subjective, experiential effect on the level of consciousness has a corresponding substrate in actual processing in the cognitive and neuronal system.

---

[12] *Extreme ironing:* 'an extreme sport which involves taking a battery powered iron and an ironing board to a remote or dangerous location and ironing some items of clothing'.

Egosurfing: 'the activity of searching the World Wide Web for occurrences of your own name'.

Shopgrifting: 'the activity of purchasing something from a shop, using it, and then returning it within a specific period in order to get a full refund'.

Speed networking: 'a method of making a potential business contact by briefly talking to a series of people at an organised event and exchanging contact details'.

Speed dating: 'a method of meeting a potential romantic partner by briefly talking to a series of individuals at an organised event, and indicating whether you are interested in seeing any of them again'.

Warchalking: 'the activity of drawing chalk symbols in public places in order to indicate the location of wireless Internet access points'.

Stagephoning: 'the activity of talking on a mobile phone in an animated and deliberately audible manner, especially in order to impress people'.

Nanopublishing: 'low-cost online publishing which uses techniques based on blogging (writing weblogs) to target a specific audience'. (All definitions taken from the Macmillan *Word-of-the-Week* website.)

6. The Mental Lexicon, Entrenchment and Network Building

By definition, new words do not have an entry in language users' mental lexicons. The general question to be addressed here is whether and how this changes in the course of the gradual establishment of new words in language users' minds and brains. More specifically, do newly acquired complex lexemes gradually get an entry of their own or is their meaning computed from that of their constituents? Is there a point or stage at which language users gain immediate and direct access to the (lexicalized) meaning of the compound and stop activating the meanings of the constituents? Given that the mental lexicon is generally held responsible for both storage of words and on-line constructions of meanings,[13] which processes does it give precedence to, processes of storage or processes of computation? And finally, is there an experimentally testable correlate to the effects of hypostatization?

6.1. Nonce-formation and pseudo-concepts

Like in other areas of language processing research, the comprehension of neologisms has been studied much more extensively than their production. Therefore, I will start out here from the perspective of the addressees of nonce-formations.

6.1.1. Comprehension

When confronted with a previously unknown complex lexeme, hearers have no choice but to rely on analytic interpretations, since a search for an entry of the whole word in their mental lexicon yields no result. Essentially, analytic understanding is based on three types of information: the meanings of the constituents, hearers' knowledge of the instantiated word-formation pattern and its known 'meanings', as well as any relevant information retrievable from the context. Hearers will activate and bring together the forms and meanings of the known constituents to form a contextually plausible and appropriate interpretation by means of a non-routinized, non-automatic process. The result of this endeavour is a mental representation which, adapting the term from Vygotsky (1962) who uses it to describe a stage in ontogenetic concept-formation, I refer to as a *pseudo-concept*.[14]

---

[13] A useful general definition of the notion of *mental lexicon* is "the cognitive system that constitutes the capacity for conscious and unconscious lexical activity" (Jarema & Libben 2007b, 2), which stresses the dynamic abilities of the system.

[14] For more details on the nature of the conceptual combination required for the understanding of novel compounds see Downing (1977), Ryder (1994), Štek-

Note that the exact conceptual status of these pseudo-concepts is far from clear. One aim of this paper is to contribute to a better understanding of this issue (see Section 7).

The amount of cognitive effort and time required for the construction of this pseudo-concept depends on a number of factors. The most uncontroversial of these is the richness of information provided by the linguistic cotext and extra-linguistic context.[15] With regard to linguistic cotext, Baayen & Neijt (1997) conducted a large-scale corpus study on the Dutch suffix *-heid*, which is similar to English *-ness* and German *-heit* in its function of forming de-adjectival abstract nouns. They showed that when texts contained hapax legomena using this suffix (which they considered novel because of their one-off character in the corpus) writers tended to relate them to the cotext, e.g. by mentioning the base or providing semantic clues in the cotext. Baayen and Neijt refer to this phenomenon as *contextual anchoring*. While this is of course evidence of speaker rather than hearer behaviour, it still supports the plausible assumption that cotext and context have a strong influence on all steps required for the construction of pseudo-concepts representing novel complex lexemes (cf. Aitchison 2003, 178).

In the sequence of steps required for processing, the first stage – if the search for an entry for the whole word fails – consists in the morphological segmentation required for understanding. If the morphological structure of the form is transparent and unambiguous, segmentation proceeds faster than in the case of multiple parsing possibilities (Aitchison 2003, 181f.). This was shown in psycholinguistic experiments carried out by Libben (1994) and Libben, Derwing & de Almeida (1999), using as stimuli fabricated compounds with ambiguous morphological structures such as *clamprod* (*clam-prod* 'an instrument for prodding clams' vs. *clamp-rod* 'a rod that is part of a clamp'). In everyday life, i.e. outside psycholinguistic or neurolinguistic experiments, structural ambiguities of this type are presumably fairly rare. One case in point could be the recent formation *shopgrifting* (see n. 12 above). A longitudinal web-as-corpus study I have carried out taking five samples over a period of six months (from November 2006 to May 2007) suggests that this word does not seem to be catching on in the internet community. The number of Google hits ranged between 38 and 50, and practically all of them were metalinguistic comments on the meaning rather than actual uses. While this could (as we would hope) be due to the limited social relevance of the term, the opacity of the constituent *grift* –

auer (2005), Gagné & Spalding (2006). In Schmid (forthcoming) I propose an account in the framework of Conceptual Blending Theory (cf. Fauconnier & Turner 2002).

[15] Despite their interest in the semantics of new words, linguists like Downing (1977), Ryder (1994) and Štekauer (2005) have deliberately opted for context-free testing in order to eliminate the influence of this unruly determinant on the comprehension of novel formations.

possibly a blend of *graft* and *lift*, the latter alluding to *shoplifting* – could play a role in the reluctance of the web-community to institutionalize the new word.

Once the new compound has been successfully parsed, the ease of pseudo-concept formation depends on the familiarity with the constituents of the compound and the semantic relation construed to exist between them. In psycholinguistic, neuro-linguistic and corpus-linguistic studies the notion of *familiarity* is usually operationalized in terms of frequency of occurrence. Since by definition nonce-formations occur for the first time, the frequency of the new word is simply not an issue. However, studies on several languages have demonstrated that the relative frequency of the constituents of a complex lexeme is a major determinant of the speed of the identification of the whole. Complex lexemes consisting of highly frequent constituents are identified and processed faster than those with rare constituents, and this suggests that it will be easier to understand novel forms consisting of frequent morphemes.[16]

Interestingly, counter to linguistic theory, which stresses the prominent role of the head constituent, there is robust evidence from these studies that the first constituent is more important for compound recognition than the second. This turns out to be true across languages irrespective of whether right-headed compounds (as in English) or left-headed ones dominate. It is possible that this is a specific manifestation of the general salience of the beginnings of words in lexical processing (Aitchison 2003, 138–140).

The ease with which novel forms are processed also correlates with the number of types of lexemes that include the constituents as components, i.e. the size of their morphological families (Schreuder & Baayen 1997, de Jong *et al*. 2002, Booij 2005). To illustrate this: it will not require much effort to process new formations using the frequent prefix *over-* (in the sense of 'excessively') as listed in the additions to the *OED* online, such as *overbill*, *overbudget* and *over-check*. On the other hand, the creative neologism *sheeple* ('people who are easily persuaded and tend to follow what other people do', Macmillan *Word-of-the-Week* website), will cause considerable difficulties, not only because of parsing uncertainties posed by a blend of *sheep* and *people* (cf. Lehrer 1996), but also because of the comparatively small cohort of established lexemes including the components *sheep* and *people*. In terms of network theories of the mental lexicon, the facilitating effect of family size can be interpreted as the result of a more densely structured network.

Familiarity not only with the constituents, especially the first one, but also with the semantic relation between them has an influence on the ease of interpretation of novel compounds (Gagné 2002, Gagné & Spalding 2006).

[16] See Taft & Forster (1976), Ahrens (1977), Sandra (1994, 108–111), Blanken (2000), Jarema (2006, 54f., 69), Semenza & Mondini (2006, 72).

This is a particularly interesting finding for the long-time student of English word-formation, since it suggests that the 'meanings' of word-formation types and patterns described in great detail in language-immanent approaches (cf. e.g. Marchand 1969, Warren 1978, Hansen *et al.* 1990) do indeed have some psychological reality. Nonce-formations coined along the meanings of productive word-formation rules stand a better chance of being understood rapidly than semantically 'irregular', creative coinings (cf. Aitchison 2003, 176f.).

An example of a semantically opaque formation likely to cause processing problems upon first encounters is the noun *she-pee* discussed above. On the one hand, the word is not unmotivated, since it is formed with a phonological allusion to the rhyming *tepee* and its constituents point to two crucial semantic aspects related to its meaning. On the other hand, however, it is not formed on the basis of any of the productive and familiar English word-formation types. To be sure, the formation does have predecessors in established compounds consisting of personal pronouns and nouns such as *he-lamb, she-goat* or *she-thief* (Marchand 1969, 75). As these examples show, however, the second constituent of these somewhat obsolete-sounding compounds invariably denotes a person, animal or occasionally plant (cf. *she-beech* denoting inferiority of timber according to Marchand). *She-pee*, on the other hand, would have to be derived from an underlying sentence such as 'place where she can pee' and thus be analyzed as an exocentric compound with a locative head.

In sum, the main factors determining the ease and speed of comprehension of novel formations are

- amount of cotextual and contextual information,
- transparency of morphological structure,
- familiarity with (i.e. essentially frequency of) the constituent morphemes, especially the first one,
- family size of the constituent morphemes, and
- familiarity with the semantic relation between the constituents.

The last three factors mentioned can be subsumed under the notion of semantic transparency, as the meaning of a new form will be transparent to hearers to the extent that its constituents and the relation between them is familiar to them.

6.1.2. Production

Concerning the actual sequence of cognitive events that can be hypothesized to take place in the mind of a speaker producing a nonce-formation, it will be useful to place the discussion in the well-known model of language production proposed by Levelt (1989) and its later refinement concerning the

role of the mental lexicon as summarized in Jescheniak (2002). The model distinguishes three modules: *conceptualizer*, *formulator* and *articulator*. Only the first two are of concern here, and I will focus on those aspects relevant for the discussion of (new) words. The *conceptualizer* formulates a preverbal message consisting of a syntactic frame and a blueprint for information structure as well as, importantly, concepts activated for the intended message. The *formulator* maps this preverbal message onto a linguistic representation, first accessing and retrieving lemmas (i.e. semantic and syntactic information of the words activated) and then, in a second step, retrieving their phonological form. Essentially, then, we can distinguish the three phases of concept activation, lexical access and lexical retrieval.

The phases relevant for the study of nonce-formations are conceptualization and access to the mental lexicon. Since the preverbal message outputted by the conceptualizer contains a syntactic frame and concepts, we have to assume that the speaker already has an idea of the type of concept to be expressed. More specifically, it should be clear at this stage whether a thing-concept, a property-concept or an event/activity-concept typically encoded by nouns, adjectives and verbs respectively is to be encoded. By definition, lexical access for nonce-formations cannot target an entry in the mental lexicon, so the most likely path will lead to the morphological, semantic and syntactic properties of the entries of the roots and affixes to be combined. Ease of access to these elements will be determined by the same factors at work in comprehension, essentially the frequency of morphemes, morphological patterns and semantic relations. What is important is that the type of concept activated will presumably prime the activation of word-class specific word-formation patterns, again depending on their frequency. I will return to this assumption further down in Section 7 below.

What are the factors determining the individual kinds of nonce-formations produced? In linguistic theorizing, this question has been dealt with extensively under the abstract label of *productivity* and has received considerable attention over the past few years.[17] While most of this body of recent work does not focus on the individual speaker using a novel form but on the availability, probability and profitability of different word-formation patterns, some insights on processing aspects can be extrapolated from these findings. Useful sources explicitly focusing on psycholinguistic aspects of new-word formation include the pioneering article by Brekle (1978) as well

[17] Even a small selection restricted to volume-sized treatments and state-of-the-art summaries has to mention Plag (1999 and 2006), Bolozky (1999) on Hebrew, Rainer (2000 and 2005), Bauer (2001 and 2005) and Dressler (2007). The exciting work of Dutch researchers, a lot of which combines psycholinguistic testing and corpus-linguistic analysis, deserves particular attention, cf. e.g. Baayen & Renouf (1996), Baayen & Neijt (1997), Schreuder & Baayen (1997), Jescheniak (2002) and Baayen (2007).

as Bauer (2001, 112–143), Jescheniak (2002), Aitchison (2003, 174–187) and Booij (2005, 231–254).

To begin with a very general observation, speakers producing novel forms – just like their hearers – favour new formations that are phonologically, morphologically and semantically transparent (Aitchison 2003, 181f.). Morphologically opaque items like those quoted in the previous section (e.g. *clamprod, shopgrifting* or *sheeple*) are not likely to be produced, unless the speaker is trying to be funny or deliberately decides to create a 'weird' and eye- (or ear-) catching form (cf. Lipka 2000).

Secondly, adult speakers show a very strong tendency in their production of nonce-formations to abide by the regular patterns described in the word-formation literature as *morphological rules* (and restrictions on them), w*ord-formations types*, *templates* (Ryder 1994) or *schemas* (Tuggy 2005; see Aitchison 2003, 174–180). The major forces behind this conformist behaviour are paradigm pressure and coercion by analogy (cf. Bauer 2001, 71–97). According to Becker (1990, 17f.), formations that are supported by larger networks of existing analogous structures are more likely to be produced than isolated ones. To some extent, this corresponds to the *family size* effect observed for comprehension in Section 6.1.1. above, and it reflects the different degrees of entrenchment in the cognitive system discussed in Section 6.2. below. While this is easily forgotten if one favours a language-immanent view in linguistics, the patterns so painstakingly analyzed and described by linguists of course only exist because they are wired into the linguistic systems in the minds of the speakers of a language.

As the literature on productivity suggests, speakers are also astonishingly adept at distinguishing analyzable but unproductive patterns from currently productive ones, and thus rarely produce new forms on the basis of no longer productive patterns. As Bauer suggests, there are "psychologically real distinctions between available ('living') and unavailable ('dead') processes" (2001, 211). Surprising as it may seem, this means that there is no direct correspondence between the ability to decompose complex lexemes (Aitchison's "back-up store", 2003, 135f.) and the ability to produce novel ones (i.e. the "lexical tool-kit", Aitchison 2003, 186f.). Language users have no difficulty in segmenting and analyzing established products of no longer productive word-formation types, for instance suffixation with *-ment*, but they still hardly ever use the patterns to form new words (cf. Bauer 1983, 55; 2001, 54, 151f.).

As already noted, the tendency of speakers to stick to the productive word-formation patterns has presumably cognitive reasons because the patterns are firmly entrenched schemas abstracted from language use (cf. Bybee 2006). But it may well have pragmatic reasons as well and be ultimately due to considerations of the principle of cooperation. After all, speakers want to make themselves understood, and their chances of getting

their message across in spite of the use of a new word are better if this word is formed on the basis of a familiar pattern.[18] However, this pragmatic aspect does not refute the idea that speakers have tacit but deeply entrenched knowledge of the word-formation rules of their language, because this is clearly a precondition of their being able to make informed guesses about the rules stored in their hearers' minds.

To be sure, 'illegal' formations are by no means unheard of and not all of them are conscious acts of creativity (cf. Bauer 2001, 62–71). Nevertheless, as the example *shopgrifting* has already suggested, it seems that irregular forms are probably less likely to catch on in the long run, because hearers are more reluctant to process them. The result on the socio-pragmatic level is that forms of this type are less likely to become institutionalized in large sections of the speech community (cf. Aitchison 2003, 182–184).[19] An interesting recent case is the noun *bouncebackability*, whose draft entry for the *OED* online dated June 2006 reads 'the capacity to recover quickly or fully from a setback, bad situation, etc.' and includes the usage label 'Chiefly *sports*'. The development of this word has been closely monitored by Hohenhaus (2006). As Hohenhaus argues, the noun may in fact turn out to be 'non-lexicalizable', because it violates the rule that the suffix *-able* can only be attached to simple transitive verbs (rather than intransitive phrasal ones).[20]

Thirdly, it is a truism that speakers often do not even notice that they have 'coined' or at least used a new word. If speakers heed the cooperative principle and try to make themselves understood, we would expect that the degree of awareness of the newness of a form is influenced by the degree of semantic transparency. Less transparent formations are more likely to reach a level of conscious processing than entirely compositional ones. Given the planned and edited nature of written texts, consciously used nonce-formations may well be more frequent there than in spoken production, but markers of newness (such as gestured inverted commas) show that awareness also occurs in oral output. If speakers become aware of producing or having produced a novel form, they will compute a) whether the new word is only new to them but might be familiar to the hearer, and, if this is not the case, b) to what extent it will be decodable for the hearer in the given context. Evidence for these processes comes from discourse itself: if speakers

[18] Of course, deliberate violations of word-formation rules and restrictions on productivity are common in the service of the attention-seeking function of neologisms in journalese, advertising copy and humorous texts or spontaneous utterances.

[19] Note that this can of course be overridden by other well-known factors contributing to institutionalization such as social relevance of the denotatum, social importance of the coiner or the humorous potential of a new word.

[20] Exceptions to this rule given by Hohenhaus (2006, 24, n. 9) are *laughable* and *knowledgable*.

decide that decoding could be difficult, they tend to select one or several of a whole range of strategies to mark the newness of the word.

These metalinguistic markers include explicit assertions of the newness of the word, e.g. by means of gambits like *what you may call* or *so-called*, the use of definitions or explanations of the meanings, inverted commas (in writing or gesturing) and, increasingly applied in computer-mediated communication, hyperlinks to definitions found elsewhere on the net (cf. Hohenhaus 1996, 139–142; Smyk-Bhattacharjee 2006, 31–34). In her study of neologisms in *The Guardian*, Fischer (1998, 176–178) found that the strategies selected correlated with the degree of institutionalization a particular new word has reached. She describes three phases:

> 1) a detailed description of the meaning of the new lexeme is offered; 2) brief paraphrasing is used; and 3) no meaning cues are offered to explain the lexeme. (Fischer 1998, 176)

The frequent strategy of adding inverted commas is illustrated in the following attestation of *fat tax* in *The Daily Telegraph*:

> A Downing Street-based policy unit has proposed a plan to place a "fat tax" on junk food in an attempt to tackle the rising incidence of heart disease [...]. (*The Daily Telegraph,* 19th February 2004; quoted from the Macmillan *Word-of-the-Week* website, s.v. *fat tax*)

Also common is the combination of *so-called* with an explanation or explanatory illustration as found in *The Guardian*'s introduction of *saviour sibling*:

> The era of the so-called saviour sibling appeared to have arrived yesterday as doctors applauded the birth of Jamie Whitaker – called into the world to allow his sick older brother Charlie to live [...]. (*The Guardian*, 20th June 2003; quoted from Macmillan *Word-of-the-week* website, s.v. *saviour sibling*)

What is important for the cognitive perspective is that the existence of these strategies indicates that writers are aware of the newness of their nonce-formations. The fact that opaque forms are more likely to be marked as new suggests that they are more likely to reach a level of conscious processing than semantically transparent ones.

Fourthly, speakers producing a novel compound are invariably forced to reduce the wealth of information they want to get across because, at least in English, they tend to restrict themselves to a rather small number of constituents, especially when they opt for compounding. While compounds or compound-like sequences like *holiday car sightseeing trip* discussed by Carstairs-McCarthy (2002, 76) are clearly not unusual, they do not seem to be favoured by most speakers (cf. Schmid 2005, 210–214).[21] Arguably, the

[21] This neglects the increasingly common but still fairly rare information-condensing phrasal nonce-compounds of the type lexicalized, e.g. in *forget-me-not*,

cognitive basis of this tendency is the need to profile certain aspects of the thing or scene in mind (Schmid 2005, 105–109). That such profiling processes indeed take place can be gleaned from a lexical error produced by one of my children. Lost for the established German word *Mundharmonika* (lit. 'mouth harmonica'), the boy unconsciously replaced it with *Blasharmonika* ('blow harmonica'). Both words suggest a scene of someone playing a wind instrument using his or her mouth. But whereas the conventional N+N compound *Mundharmonika* highlights the location of the activity while backgrounding the activity itself, the complementary pattern can be observed for the V+N compound *Blasharmonika*.

Similar profiling processes (resulting in underdeterminacy) presumably take place when new words are produced. Let us assume that, for some reason or other, a speaker is eager to use one single word to refer to the feeling of relaxed and content laziness and languor so common after a substantial and late Sunday morning breakfast. If the phrasal-compound pattern of the type *Sunday-morning-after-breakfast-laziness* is not strongly entrenched in the speaker's mind, the options left include compounds such as *breakfast-langour* or *Sunday-morning off-time*, which invariably omit references to crucial aspects of the feeling to be expressed. (Note that even the phrasal compound leaves out aspects such as 'substantial').

These are by no means far-fetched illustrations. Authentic examples are the compounds *noise pollution* mentioned in Section 3 above, and *fat tax* (see Section 6.2.1. below), which is explained on the Macmillan *Word-of-the-Week* website as 'a tax on foods which are considered to be unhealthy, especially fatty or sweet foods which can lead to obesity or other health problems'. Like my fabricated example, this N+N compound profiles only two aspects for attention but does not include explicit morphological clues referring to other elements included in its meaning such as 'unhealthy', 'sweet' or 'can lead to obesity'. Instead it relies on the notion of *fat* standing metonymically for 'unhealthy' and 'cause of obesity'.

Thus, while being relatively motivated and therefore often at least apparently transparent, English compounds inevitably underdetermine the meaning intended, let alone the rich conceptual content potentially associated with them by the speakers in given nonce-formation situations. Plausibly, this is one cognitive source of the astonishingly high rate of instantaneous coinings (see Section 3 above).[22] The idea that compounds invariably

what-you-may-call-it, whodunnit. Attested nonce-formations quoted by Hohenhaus (1996, 349f.) include *reporter-turned-hostess-turned-novelist* and *split-the-difference deal*.

[22] The second cognitive source of instant idiomatization of compounds can be found in the fact that two concepts are conflated in one new concept. That such a merging of concepts invariably results in extra emergent meaning was already noted by de Saussure ([1916] 1983, 130f.) and is of course a major insight of Gestalt psychology (cf. Schmid forthcoming).

rely on pragmatic, extra-linguistic information was already emphasized by Downing (1977) and Bauer (1979), but has gained much more momentum recently. It is not only supported by linguists taking a cognitive approach to word-formation (cf. Ungerer 2002; Schmid 2005, 106; Štekauer 2005, 249–251; Tuggy 2005, 238f.), but recently also by Aronoff (2007, 61–68), who has apparently changed his mind on the question of the compositionality of products of word-formation.

A related point is more speculative but of considerable theoretical interest. If speakers are indeed predisposed in their coinings by the frequent and productive word-formation patterns available in their language, this may in fact lead them to highlight certain aspects of the scenes they have in mind when producing a novel form, while backgrounding others. For example, speakers of Italian could be more prone to produce a novel V+N-compound than speakers of English, because this pattern is much more frequent in the lexicon of Italian – and thus entrenched in their minds – than in English: compare, for instance, the Italian V+N compounds *portarifuti*, lit. 'carry rubbish', and *salvadanaio*, lit. 'save money', to their English N+N equivalents *dustbin* and *moneybox*, respectively. Whether the dynamic aspects highlighted by the verbs will also be more prominent in their minds because of this linguistic predisposition is an interesting Whorfian question deserving future work.

6.2. On the way to entrenchment

6.2.1. Frequency of exposure

Linguists, psycholinguists and neurolinguists more or less agree that even a single exposure to a word has some sort of effect on the cognitive system. A lucid account of this position is provided by Sandra (1994, 30):

> The hypothesis proposed here is that the mere occurrence of a word will inevitably lead to a pattern of activation in a memory substrate. If the word has no representation yet, this pattern will automatically leave a representation in memory, although a very fragile one.

The idea that even a single encounter with a new word leaves a trace in memory has recently been confirmed experimentally by de Vaan, Schreuder & Baayen (2007).[23] With regard to repeated exposure, Sandra (1994, 30f.) goes on to state that

> upon subsequent encounters with the word the same representation will be contacted, each contact resulting in the strengthening of the representation. The

[23] Interestingly, there is experimental evidence suggesting that the establishment of a lexical representation of a new word is facilitated by sleep. Dumay and

> stronger the representation becomes the easier it will be accessible from a stimulus, a frequency effect which has been widely attested in the word recognition literature.

In fact, the frequency effect is, as Knobel, Finkbeiner & Caramazza (forthcoming) recently put it, "perhaps the most robust effect in all of psychology". We should add to this that if repeated exposure does not occur, the memory trace will become weaker and eventually fade away, i. e. the word is forgotten. If this happens in the minds of all initial users and hearers of the word, and the word is not recorded in written language, it has no chance of becoming institutionalized (Aitchison 2003, 174). Particularly good candidates for such a fate are nonce-formations of the type *the/this X thing/ business* dubbed "dummy-compounds" by Hohenhaus (1996, 281–288 and 1999; examples quoted by Hohenhaus include *the dress code thing* and *the degree business*).

Two early and much-quoted statements of the relation between entrenchment and frequency can be found in Bybee (1985) and Langacker (1987).[24] These prepared the ground for the current craze about usage-based approaches in grammaticalization theory, cognitive linguistics and language acquisition. Bybee (1985, 117) wrote that

> if we metaphorically assume that a word can be written into the lexicon [i.e. the mental one, HJS], then each time a word in processing is mapped onto its lexical representation it is as though the representation was traced over again, etching it with deeper and darker lines each time. Each time a word is heard and produced it leaves a slight trace in the lexicon, it increases in lexical strength.

Langacker describes the gradual entry of a new complex lexeme in the mental lexicon by means of two metaphors which highlight different aspects. The first relies on the notion of *entrenchment*, which is increasingly used not only in cognitive linguistics (cf. Schmid 2007) but also in grammaticalization theory (cf. e. g. Croft 2000, 32, 72f.; Krug 2003, 15; Hoff-

---

Gaskell (2007) found that words encountered at 8 p.m. did not induce symptoms of establishment immediately, but did so after a 12-hour period including a night's sleep, and continued to do so after 24 hours. Words learned at 8 a.m. did not show such effects immediately or after 12 hours of wakefulness, but showed them after 24 hours, after sleep had occurred.

[24] It should not go unnoticed that similar ideas on the effect of recurrent use on the degree of entrenchment were already put forward by John Locke in his *Essay Concerning Human Understanding:* "There comes, by constant use, to be such a connexion between certain sounds and the ideas they stand for, that the names heard, almost as readily excite certain ideas as if the objects themselves, which are apt to produce them, did actually affect the senses" (Locke [1698] 1975, 3.2.2).

mann 2004, 179, 189–194; see also Bybee 2006). According to Langacker (1987, 59), there is a

> continuous scale of entrenchment in cognitive organization. Every use of a structure has a positive impact on its degree of entrenchment, whereas extended periods of disuse have a negative impact. With repeated use, a novel structure becomes progressively entrenched, to the point of becoming a unit; moreover, units are variably entrenched depending on the frequency of their occurrence.

The source domain of the second metaphor is the scaffolding of a new house or building. In this image the morphological constituents making up a novel complex lexeme correspond to the scaffold, which facilitates the construction of the house but can be taken away when the structure is in place and is gradually completed. This metaphor highlights the fact that the meaning of entrenched composite units do not tend to be compositional any more.

Langacker's early account leaves two important questions unanswered. The first one concerns the precise nature of the relation between frequency and entrenchment. For example, what is it that we have to measure the frequency of? And do we have to imagine that there is a simple linear correlation between frequency and entrenchment, in such a way that a word that is three times more frequent than another is assumed to be three times as deeply entrenched (whatever that means)?

What has been said about this relation in the previous section suggests that not only the frequency of the complex form fosters entrenchment but also the frequency of the constituents. This should be taken into consideration. With regard to frequency of words in general, both simple and complex, the effect of frequency has commonly been modelled in the form of a logarithmic function.[25] This is at least the pattern suggested by reaction times in naming experiments, where "at the low-frequency end of the spectrum, small changes in frequency have very large impact on reaction time, but at the high-frequency end, large changes in frequency have negligible effects" (Forster 2007, 42). Interestingly, a non-linear development was also found for the institutionalization of neologisms in the corpus study of newspapers by Fischer (1998, 174). Transferred to the entrenchment of complex lexemes this means that the pace in the increase of entrenchment will be faster during the first repetitions of exposure to a word than later, which confirms the intuitive impression that very few occasions of processing a new word suffice to create the feeling of having acquired a concept associated with it (see Section 7 below).

The second question raised by Langacker's two metaphors is how frequency and degree of entrenchment, on the one hand, and semantic transparency and (lack of) compositionality, on the other, interact. The scaffolding metaphor essentially suggests that both the strength of entrenchment

[25] See Forster (2007, 42–44) for other approaches, e.g. a linear function of the frequency rank (rather than frequency score).

(i. e. storage in the mental lexicon) and the degree of opacity are more or less a function of frequency of usage. On the other hand, Langacker also notes that "the familiarity of a complex expression does not blind us to its componentiality and render us unable to perceive the contribution of individual components" (1987, 461). In assessing this relation it is important to keep in mind the observation that even in their initial uses as nonce-formations novel complex lexemes tend to be much less compositional than is commonly believed.

6.2.2. Semantic transparency, storage and processing

The role of semantic transparency in the storage and processing of compounds has been on the agenda of psycholinguists and neurolinguists for quite a while but has recently received particular attention. In principle, there are three idealized types of competing theories – none of which, incidentally, addresses the question of how semantic transparency arises in the first place. Firstly, *non-decompositional*, *direct-access* or *whole-word* models assume that complex lexemes are stored holistically and accessed directly just like simple lexemes, irrespective of their degree of compositionality. A typical representative of this type is the model derived by Butterworth (1983) from the analysis of speech errors and the speech production of aphasics. Quite obviously, non-compositional models run into difficulties when it comes to explaining the indubitable decomposition abilities that speakers need and readily demonstrate in understanding novel complex lexemes (cf. Aitchison 2003, 135f., 186f.). Secondly, *compositional* (or *decompositional*) theories see the morpheme as the standard unit of lexical access and opt for a computational processing of complex lexemes (e. g. Taft & Forster's (1976) so-called *prefix stripping* model). Models of this type reduce the amount of storage capacity required at the cost of computational effort. They turn out to be unconvincing for highly opaque word-formation items as well as for highly frequent ones, since both types would be processed more rapidly and efficiently if they could be accessed wholesale rather than having to be computed anew each time. A third type of models, *dual-access* or *race-horse* models, argues for a competition between morphemic and whole-word access (cf. e. g. Frauenfelder & Schreuder 1992, and Caramazza, Laudanna & Romani 1988 on inflectional morphology). They assume that while one access routine searches for an entry matching the complex lexeme, the other one seeks contact with the addresses of the constituent morphemes. Models of this type have been criticized for their lack of efficiency caused by the wealth of redundancies generated, and for their inability to explain the apparent links between the whole-word representation and the representation of the constituents (Libben 2006). In spite of this, dual-access models are currently favoured by many researchers,

among them Baayen & Neijt (1997, 584–586), Aitchison (2003, 135f.; 2005, 1785f.), Booij (2005, 236), Jarema (2006) and Libben (2006).

A good starting-point for reviewing some of the evidence for these models is Sandra's (1990) priming study on the effect of semantic transparency. In priming tasks, a *prime* or context stimulus precedes a target stimulus. Response latencies, i.e. reaction times, are measured for tasks like verification or lexical recognition of the target stimulus. The reasoning behind this test design is that contextually relevant primes (e.g. *bird* as a prime for the target stimulus *robin*) will speed up response latencies because they facilitate the activation of information needed for the target task. In this experiment, response latencies in lexical decision tasks were measured using opaque and transparent English compounds as stimuli, and primes targeting the meanings of individual constituents. For example, the first element of the transparent compound *birthday* was successfully primed by its antonym *death*, the first element of the opaque compound *Sunday* – perhaps not an ideal choice – was primed by *moon*. Priming effects were only obtained for transparent but not for opaque compounds. This led Sandra to two conclusions: transparent compounds are not accessed directly but via a decomposition into their constituent morphemes, while opaque compounds are accessed directly (since they showed no constituent priming effects) and thus most likely stored holistically in the mental lexicon. Although similar evidence was obtained for Dutch compounds by Zwitserlood (1994), more recent research has not fully confirmed this finding.

A study by Libben *et al.* (2003, co-authored by Sandra) produced support for a race-horse type of model. Compounds of four types of semantic transparency were used in this study: fully transparent ones such as *bedroom*; compounds with an opaque first and a transparent second element (*strawberry*); the counterpart with a transparent first and an opaque second element (e.g. *jailbird*); and fully opaque compounds such as *hogwash*, in which neither constituent had a relation to the holistic meaning. The result of a primed lexical decision task was that both initial and final constituents prime all compound types, i.e. fully compositional, partly compositional and non-compositional ones. Since, however, it was also found that compounds with opaque heads take longer to process, Libben *et al.* concluded that both types of access seem to be possible.

Findings of this type have led Libben to argue for a modified version of a race-horse model based on a principle dubbed "maximization of opportunity" (2006, 12 and *passim*). This principle explicitly contests the long-standing belief that the mental lexicon in fact strives for economy, and "posits a mental architecture in which all representations that can be activated will be activated" (Libben 2006, 12). Also in contrast to widespread assumptions, Libben (2006, 14) cites experimental evidence suggesting that, at least initially, opacity does not actually diminish constituent activation. The response latencies found in Libben *et al.* (2003) are explained by an inhibitory process that is caused by a mismatch between the semantic rep-

resentation of the opaque compound and that of the constituents. The advantage of this model over earlier race-horse models is that it accounts for the links between the representation of the whole word and the constituents by including the matching procedure (which requires extra time during the processing of opaque compounds).

6.2.3. Frequency and semantic transparency

To take stock of the discussion so far, in Langackerian terms, the scaffolding often stays in place even when the construction work is completed and the building has taken shape. The probability of storage as opposed to computation seems to be a function of both frequency and semantic transparency. If we cross-tabulate the effects of these two factors, the following picture emerges:

- There seems to be little doubt that frequent opaque complex lexemes do get an entry of their own in the mental lexicon (cf. e.g. Baayen & Neijt 1997, 568). However, access to their constituents is apparently by no means blocked, not even when neither of the constituents has a link to the composite meaning. Yet again, activation is inhibited during such an early phase of lexical access that the language user does not even become aware of it.
- Most likely, frequent transparent complex lexemes – which are probably much rarer than is usually assumed – are also entrenched. They are stored in the mental lexicon and accessed directly (cf. Bauer 2001, 122), but their constituents stand a better chance of easily reaching the level of conscious processing. While these complex words may not be idiomatized, they are still conventionalized ways of referring to things and ideas (Sauer 1992, 117; Booij 2005, 235). The type of processing preferred also depends on the frequency of the constituent morphemes (cf. Hay 2001). If the base is significantly less frequent than the derived form, the latter has a good chance of being stored. For example, since the prefixation *insane* occurs much more frequently than its base *sane* it is very likely that the complex form is stored and accessed directly (Hay 2001, 1044).
- The evidence reported suggests that rare transparent complex lexemes will tend to be processed computationally, since there is little to be gained by holistic entrenchment (Baayen 1993, 181). However, taking into account the logarithmic form of the frequency effect, we can predict that after the first few encounters holistic processing will soon gain the upper hand. What is more, what may be a rare word for one speaker may well be already familiar to another (see Section 7 below).
- Rare opaque complex lexemes will presumably be the best candidates for a comparatively time-consuming race-horse type competition between

> the search for a holistic entry and computational processing of the constituents.

For linguists, who tend to be more interested in meaning, structure and use of language than in the finer details of lexical access, post-access decisions and retrieval procedures carried out in fractions of seconds, the preceding review leaves a number of questions unanswered regarding the relation between frequency and semantic transparency.

Firstly, it is unclear whether frequent exposure is a necessary condition for the emergence of opacity. As I have already pointed out, this is probably not the case, since even many nonce-formations are not semantically transparent outside their original context (see Section 3 above). In fact, it does not seem implausible that instantaneous coinings with partly or fully non-compositional meanings (e.g. metaphorical or metonymic compounds) actually need a smaller number of repetitions for entrenchment than fully transparent ones, because their interpretation requires a greater processing depth which yields stronger traces in the cognitive system. Due to their enormous potential for attracting attention, semantically and/or morphologically irregular forms that strike the hearer as being funny or particularly creative and look as if they are indeed unlexicalizable could well be the best candidates for achieving almost immediate entrenchment. This is exploited in advertising discourse by the invention of eye-catching formations like German *unkaputtbar*, lit. *un-broken-able* 'impossible to break' (Hohenhaus 2005, 369).

Secondly, and in a sense complementary to the first question, it has been a matter of debate whether frequency fosters not just entrenchment but also opacity. In other words, are frequent complex lexemes more likely to be opaque than rare ones, or do neologisms automatically idiomatize if they catch on? Aronoff (1983, 168) argues on the basis of a frequency count of words ending in *-iveness* and *-ivity* in the Brown corpus that there is indeed a causal link between semantic complexity and token frequency. However, according to Bauer (2001, 50f.), the evidence is not unequivocal but open to alternative interpretations. Hay (2001, 1066), in the study already mentioned above, concludes that

> a low-frequency form is likely to be nontransparent if it is composed of even-lower-frequency parts. And a high-frequency form may be highly decomposable if the base it contains is higher frequency [sic!] still.

Why should more frequent words be more prone to idiomatization than rare ones in the first place? The answer to this question may be revealed by a look at the cognitive consequences of repeated exposure to a new or recent word. This repeated exposure does not simply result in increasing entrenchment in memory, as claimed by Sandra (1990) and Langacker (1987) as cited above and in many other sources. Rather, diverse occurrences of the same word in different contexts lead to a constant enrichment of the set of conceptual

associations, or the cognitive model (cf. Ungerer & Schmid 2006, 47–58) automatically activated by it. With more and more repetitions the word gets increasingly loaded with conceptual content. In network models, this development is described as an increase in density and multiplexity in the conceptual (and neurological) network that is the mental lexicon.

To resort again to a rather personal example, I can still recall being confused by the then new term *computer virus* when I first came across it in the late 1990s, because I took the metaphorical reference to bugs too literally. After repeated exposure to the expression in various written and spoken contexts, my misguided pseudo-concept not only gradually turned into a more or less well-defined proper concept but became more and more complex and rich. I learned, for example, who the creators of computer viruses were, how the bugs found their way to people's computers, what kind of damage they could do, what types could be distinguished, and many other pieces of knowledge that are very closely associated with the concept in my mind and thus, at least in a cognitive-semantic definition of the term, part of the *meaning*. This enrichment both constitutes a growing degree of opacity in itself and has the potential to increase it still further. That a similar process takes place for initially transparent compounds is suggested by attribute listing tasks carried out by Ungerer und Schmid (1998), which revealed the astonishing amount of extra conceptual content associated with seemingly compositional compounds like *apple juice*, *kitchen table* or *coat collar*.

In short, it is not only, or at least not so much, the sheer token frequency of a new word that decides on the degree of opacity, but the diversity of contexts of usage in which a word is encountered. This diversity in turn is of course co-determined by token frequency. Support for this idea comes from a recent study by Adelman *et al.* (2006), who show that contextual diversity is a better predictor of latency times in lexical decision tasks than word frequency.

The third question that tends to be neglected in the models discussed above but is of course of major relevance to the practising linguist concerns the role of the semantic relation between the constituents and its match with the known meanings of the word-formation pattern instantiated. As already mentioned in Section 6.1.1. above, studies reported by Gagné & Spalding (2006) have shown that the familiarity with the relation expressed by a novel compound facilitates its processing by aphasic patients. This familiarity is apparently determined mainly by knowledge about how the modifier is used in other combinations (Gagné & Spalding 2006, 151f., 159f.). For example, if hearers come across a new compound beginning with *mountain*, they will bring to bear on their interpretation the expectation that it has a locative relation to the head, because they know from their experience with other compounds modified by *mountain* (e.g. *mountain cabin*, *mountain resort*) that this is the most common relation. If this expectation yields a plausible interpretation, processing is facilitated.

7. Hypostatization and Psycholinguistic Evidence: The Nature of Pseudo-concepts

Section 3 above concluded with the claim that hypostatization is at work whenever content words are used, no matter whether they are established or new. If entrenchment and storage of words in the mental lexicon are considered to be the psychological correlate of hypostatization, then we are facing a dilemma, because rare words, and thus also new ones, not surprisingly show few signs of entrenchment. To reconcile these clashing claims it is worth taking a closer look at some of the evidence from psycholinguistic and corpus-based studies.

Firstly, with regard to frequency, attention has to be drawn to the fact that psycholinguistic and corpus-linguistic studies inevitably investigate collectives rather than individuals. Experiments aim at insights that are representative of larger populations, and frequencies counted in corpora are no more than indicators of observed frequencies in one sample of texts, no matter how large it may be. While frequency of occurrence is a text-related measure whose significance is transferred to a collective, the availability of concepts is very much an individual thing. Words that are very rare in a given corpus may still be encountered very often by some speakers, who, as a result, have a very clear and deeply entrenched concept associated with the words. With regard to institutionalization, it is of course a very old insight that words can be specific to certain groups of users and certain registers. In short, 'rare' words can be familiar to and entrenched for some speakers, while being simply unknown to others.

Secondly, large portions of the psycholinguistic and corpus-linguistic literature on word-formation and the productivity of individual patterns treat these patterns as if they were very similar in their cognitive functions. For example, Baayen & Renouf (1996) investigate five de-adjectival affixes, the suffixes *-ly*, *-ness*, *-ity* and the prefixes *un-* and *in-*. To be sure, they do not lump the findings on these affixes together but give a differentiated view of the number of established formations in their corpus and the number of hapax legomena, and they emphasize that "word formation is conceptually driven" (1996, 90). What they fail to note, however, is that prefixation and suffixation have entirely different effects on their bases. While prefixation predominantly affects the conceptual content of the base, in most cases generating a contrast (Schmid 2005, 162–165), the major function of suffixation is to change the conceptual status of the information expressed by the base, i. e. to re-categorize the cognitive unit. This was already stressed by Kastovsky (1986, 595) and has recently been emphasized in cognitive approaches favouring an onomasiological perspective (Štekauer 1998, Ungerer 2007). What is more, and what is particularly important in the context of hypostatization, suffixations resulting in nouns, in fact mainly abstract nouns, turn out to be most frequent (Schmid 2005, 183–186). This suggests that the potential of suffixation tends to be exploited frequently for the

purpose of hypostatization and reification. The second most frequent target of suffixation is the word-class of adjectives, while verbs, which probably have the least strong hypostatization potential, are mainly produced by conversion.

Finally, having a fully-fledged concept of a cognitive unit as a result of repeated exposure to the word and increasing entrenchment may be one way of thinking about hypostatization in processing terms. Another could be to take more seriously the warnings of Bolinger and others that words can serve as signposts to what Bolinger calls "pseudo-entities" (1980, 61–65). It may well be that the recognition that a linguistic entity is supposed to be a word, or that it is presented by the speaker in a way that suggests that he or she thinks it is a word, is sufficient to raise the expectation in the hearer that there must be a concept or thing behind it.

This means that the conceptual effect of the use of one single word, rather than a syntactic phrase, can of course be exploited pragmatically. The use of a (novel) lexeme tacitly implies, as Bolinger puts it in the passage quoted above, "that the world puts [entities] like this in a class by themselves". What is crucial here is Bolinger's reference to "the world", because this highlights the social nature of the use of (new) words. When the authors of science fiction or fantasy literature 'invent' new props or social practises, they not only insinuate the existence of the things or activities but suggest that the members of the culture in which their narratives are set have social knowledge of the existence of the categories of things or activities. Likewise, when we come upon novel lexemes such as *shopgrifting*, we immediately form two assumptions: first, that there are people, in whatever section of society, who engage in this activity so regularly that it represents a name-worthy category; and second, that there are people, again somewhere out there, who are so familiar with this type of activity that they find it worth encapsulating in a new word.

What would a sound and efficient reaction of the cognitive system be like? It would not seem unlikely that a hearer being confronted with a new word in a specific context and automatically forming these two assumptions would embark on processing by setting up something like an address or a dummy slot for an entry in his or her mental lexicon. This could be reasonable and possible even if the first context did not include enough clues as to what the word could possibly mean. Subjectively, this could be a state where the word's form (and presumably word-class) leaves its first traces accompanied by the information that it corresponds to a concept, the details of which may not yet be clear to the language user but can be filled in the case of repeated exposure in diverse contexts. This would mean that, metaphorically speaking, the very first and fine lines etched into memory as a result of the first encounter with a word (cf. de Vann, Schreuder & Baayen 2007; see Section 6.2.1. above) would represent the impression that there is some sort of societally relevant concept out there which might turn out to be worthy of more solid entrenchment in the future. This in fact could be a

definition of the notion of pseudo-concept introduced in Section 6.1.1. above.

Note that this argument for the early arrangement of a pseudo-concept stands in stark contrast to various warnings in the literature that not all new words are used in order to create a concept. Caveats of this type were emphasized by Downing (1977), who argues that nonce-formations like *apple-juice seat* are used as "deictic devices" in purely referring function: "The use of these forms hardly seems to imply the existence of a name-worthy CATEGORY. At most, they can be said to represent a momentarily name-worthy ENTITY" (Downing 1977, 823; original emphasis). A second type of non-naming and thus non-concept-forming creation has been credited with syntactic, text-deictic or even pro-nominal functions.[26] Lipka, for example, discusses the following example, claiming that the agent-nominalization *puzzler* fulfils merely a textual function:

> Not since […] 1941 when Rudolf Hess flew off from Berlin to Scotland […] had a private trip abroad by a German leader so puzzled his countrymen. This time the puzzler was none other than Franz Josef Strauss […]. (Lipka 1987, 64; original emphasis)

While the textual function of cases of this type cannot be contested, I still claim that the pretence of hypostatization is kept up. No doubt, complex lexemes with deictic and textual functions are firmly anchored in the extralinguistic context or linguistic cotext respectively, and serve mainly a referring function; nevertheless, I believe that the use of one single word at least insinuates the existence of a corresponding category of referents. One piece of evidence for this belief is that highly context-dependent nonce-formations are often felt to have an ironic or humorous tone, which, at least partly, derives from this *categorical presupposition*.[27] It is clearly no coincidence that Lipka (1987, 65) refers to texts that play with language, ironic texts, aggressive texts like parodies and of course advertising copy as specifically rich sources of nonce-compounds. The only clear exception, where

---

[26] On the syntactic function of word-formation, see Hansen (1999, 85–88 and 2002, 195f.). The textual function of complex lexemes has recently been stressed by Baayen & Neijt (1997) and Baayen & Renouf (1996, 93); previously, it was already commented on and richly illustrated by Lipka (1977, 161f.; 1981, 129f. and 1987), Kastovsky (1978, 362f.; 1982, 165, 217 and 1986), Dederding (1983, 49–51) and Hohenhaus (1996, 255–272).

[27] Note that the existence of a category of things is not asserted but indeed presupposed. Categorical presuppositions are in fact resistant to negation. For example, the presupposition for the quote from Lipka above would be: 'There is a category of people called *puzzlers*'. This would not be cancelled by negating "The puzzler was none other than Franz Josef Strauss" to "The puzzler was not none other than Franz Josef Strauss" or "It is not true that the puzzler was none other than Franz Josef Strauss".

hypostatization does not seem to come into play at all, are *dummy-compounds* with general head-nouns like *the dress code thing* and *the degree business* quoted from Hohenhaus (1996, 281–288 and 1999) above.

In sum, it seems feasible that pseudo-concepts resulting from hypostatization are created, or at least insinuated, even by the use of nonce-formations. In processing terms, this would have to be modelled as an effect independent of the increase in familiarity and of the entrenchment brought about by repeated exposure. One way of thinking about it would be the analogy of an accommodation address that is generated automatically once a word, especially a noun, is encountered. In linguistic terms, this effect must be relegated to the level of parole rather than norm or system, even though it occurs invariably and systematically and has a strong pragmatic potential.

8. Conclusion

This paper has attempted to provide a conspectus of what happens in the minds and cognitive systems of speakers and hearers while processing novel complex lexemes and recently coined neologisms. The main factors influencing the processing and storage of neologisms and their interactions were reviewed: context; familiarity with the constituents; familiarity with the underlying semantic relation; increasing familiarity with and entrenchment of the composite form and its meaning as a result of frequent exposure; phonological, morphological and semantic transparency. The processing perspective was complemented by a more traditional concept-related perspective focusing on the hypostatizing potential of words in general and new words in particular. It was claimed that repeated exposure is not a prerequisite for the feeling experienced by the addressees of novel forms (and nonce-formations with mainly deictic or textual functions) that there is a cognitive category of entities denoted by a word. Furthermore, it was emphasized that semantic opacity does not have to be the result of a gradual process whose pace and intensity are determined by frequency of occurrence, but may arise during the act of creating a new word because of the need to profile a limited number of aspects of the scene envisaged for encoding.

Works Cited

Adelman, James S., Gordon D.A. Brown & José F. Quesada. 2006. "Contextual diversity not word frequency determines word naming and lexical decision times". *Psychological Science* 17: 814–823.

Ahrens, R. 1977. "Wortfindungsstörungen für zusammengesetzte Worte (Nomina composita) bei Aphasien". *Archiv für Psychiatrie und Nervenkrankheiten* 224: 73–87.

Aitchison, Jean. 2003. *Words in the Mind: An Introduction to the Mental Lexicon*, 3rd ed. Oxford: Blackwell.

Aitchison, Jean. 2005. "Speech production and perception". *Morphology: An International Handbook on Inflection and Word-formation.* Vol. II. Ed. Geert Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas. Berlin/New York: de Gruyter. 1778–1788.

Aronoff, Mark. 1983. "Potential words, actual words, productivity and frequency". *Proceedings of the 13th International Congress of Linguists. Tokyo, August 29–September 4, 1982.* Ed. Shiro Hattori & Kazuko Inoue. Tokyo: Gakushuin UP. 163–171.

Aronoff, Mark. 2007. "Language: Between words and grammar". In: Jarema & Libben 2007a. 55–79.

Baayen, R. Harald. 1993. "On frequency, transparency and productivity". *Yearbook of Morphology 1992.* Ed. Geert Booij & Jaap van Marle. Dordrecht: Kluwer. 181–208.

Baayen, R. Harald. 2007. "Storage and computation in the mental lexicon". In: Jarema & Libben 2007a. 81–104.

Baayen, R. Harald & Anneke Neijt. 1997. "Productivity in context: A case study of a Dutch suffix". *Linguistics* 35: 565–587.

Baayen, R. Harald & Antoinette Renouf. 1996. "Chronicling the Times: Productive lexical innovations in an English newspaper". *Language* 72: 69–96.

Barnhart, Clarence L., Sol Steinmetz & Robert K. Barnhart. 1980. *The Second Barnhart Dictionary of New English.* New York: Longman.

Bauer, Laurie. 1979. "On the need for pragmatics in the study of nominal compounding". *Journal of Pragmatics* 3: 45–50.

Bauer, Laurie. 1983. *English Word-formation.* Cambridge: Cambridge UP.

Bauer, Laurie. 2001. *Morphological Productivity.* Cambridge: Cambridge UP.

Bauer, Laurie. 2005. "Productivity: Theories". In: Štekauer & Lieber 2005. 315–334.

Bauer, Laurie & Rodney Huddleston. 2002. "Lexical Word-formation". *The Cambridge Grammar of the English Language.* Ed. Rodney Huddleston & Geoffrey K. Pullum. Cambridge: Cambridge UP.

Becker, Thomas. 1990. *Analogie und morphologische Theorie.* München: Fink.

Blanken, Gerhard. 2000. "The production of nominal compounds in aphasia". *Brain and Language* 74: 84–102.

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

Bolinger, Dwight. 1980. *Language: The Loaded Weapon.* London etc.: Longman.

Bolozky, Shmuel. 1999. *Measuring Productivity in Word Formation.* Leiden: Brill.

Brekle, Herbert E. 1978. "Reflections on the conditions for the coining and understanding of nominal compounds". *Proceedings of the 12th International Congress of Linguists. Vienna, August 28–September 2, 1977.* Ed. Wolfgang U. Dressler & W. Meid. Innsbruck: Universitätsverlag Innsbruck. 68–77.

Butterworth, Brian. 1983. "Lexical representation". *Language Production.* Vol. II. Ed. Brian Butterworth. San Diego, CA: Academic P. 257–294.

Bybee, Joan. 1985. *Morphology: A Study in the Relation Between Meaning and Form.* Amsterdam: Benjamins.

Bybee, Joan. 2006. "From usage to grammar: The mind's response to repetition". *Language* 82: 711–733.

Caramazza, Alfonso, Alessandra Laudanna & Cristina Romani. 1988. "Lexical access and inflectional morphology". *Cognition* 28: 297–332.

Carstairs-McCarthy, Andrew. 2002. *An Introduction to English Morphology: Words and their Structure*. Edinburgh: Edinburgh UP.

Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*, Harlow etc.: Longman.

Dederding, Hans-Martin. 1983. "Wortbildung und Text. Zur Textfunktion (TF) von Nominalkomposita (NK)". *Zeitschrift für germanistische Linguistik* 11: 49–64.

de Jong, Nivja H.M., Laurie B. Feldman, Robert Schreuder, Matthew Pastizzo & 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.

de Vaan, Laura, Robert Schreuder & R. Harald Baayen. 2007. "Regular morphologically complex neologisms leave detectable traces in the mental lexicon". *The Mental Lexicon* 2: 1–23.

Delazer, Margarete & Carlo Semenza. 1998. "The processing of compound words". *Brain and Language* 61: 54–62.

Downing, Pamela. 1977. "On the creation and use of English compound nouns". *Language* 53: 810–842.

Dressler, Wolfgang. 2007. "Productivity in word formation". In: Jarema & Libben 2007a. 159–183.

Dumay, Nicolas & M. Gareth Gaskell. 2007. "Sleep-associated changes in the mental representations of spoken words". *Psychological Science* 18: 35–39.

Faiß, Klaus. 1978. *Verdunkelte Compounds im Englischen: Ein Beitrag zu Theorie und Praxis der Wortbildung*. Tübingen: Narr.

Fauconnier, Gilles & Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Fischer, Roswitha. 1998. *Lexical Change in Present-day English: A Corpus-based Study of the Motivation, Institutionalization, and Productivity of Creative Neologisms*. Tübingen: Narr.

Forster, Kenneth I. 2007. "Visual word recognition: Problems and issues". In: Jarema & Libben 2007a. 31–53.

Frauenfelder, Uli H. & Robert Schreuder. 1992. "Constraining psycholinguistic models of morphological processing and representation: The role of productivity". *Yearbook of Morphology 1991*. Ed. Geert Booij & Jaap van Marle. Dordrecht: Kluwer. 165–183.

Gagné, Christina L. 2002. "Lexical and relational influences on the processing of novel compounds". *Brain and Language* 81: 723–735.

Gagné, Christina L. & Thomas L. Spalding. 2006. "Conceptual combination: Implications for the mental lexicon". In Libben & Jarema 2006. 145–168.

Gardiner, Alan H. 1932. *The Theory of Speech and Language*. Oxford: Clarendon.

Green, Jonathan. 1991. *Neologisms – New words since 1960*. London: Bloomsbury.

Handl, Sandra. 1999. "Das englische Lexikon: Motivation, Lexikalisierung und Institutionalisierung". Unpublished M.A. thesis, Ludwig-Maximilians-Universität München.

Hansen, Barbara, Klaus Hansen, Albrecht Neubert & Manfred Schentke. 1990. *Englische Lexikologie*. 2. Aufl. Leipzig: Enzyklopädie.

Hansen, Klaus. 1999. "The treatment of word-formations and word-formation patterns in a monolingual English dictionary (with special reference to the *Longman Dictionary of Contemporary English*)". *Words, Lexemes, Concepts – Approaches to the Lexicon: Studies in Honour of Leonhard Lipka*. Ed. Wolfgang Falkner & Hans-Jörg Schmid. Tübingen: Narr. 85–98.

Hansen, Klaus. 2002. "The functions and meanings of word-formations (with reference to Modern English)". *Language and Function. To the Memory of Jan Firbas*. Ed. Josef Hladký. Amsterdam/Philadelphia: Benjamins. 195–210.

Hay, Jennifer. 2001. "Lexical frequency in morphology: is everything relative?". *Linguistics* 39: 1041–1070.

Hoffmann, Sebastian. 2004. "Are low-frequency complex prepositions grammaticalized? On the limits of corpus data – and the importance of intuition". *Corpus Approaches to Grammaticalization in English*. Ed. Hans Lindquist & Christian Mair. Amsterdam/Philadelphia: Benjamins. 171–210.

Hohenhaus, Peter. 1996. *Ad-hoc-Wortbildung. Terminologie, Typologie und Theorie kreativer Wortbildung im Englischen.* Frankfurt a.M.: Lang.

Hohenhaus, Peter. 1999. "An overlooked type of word-formation: Dummy-compounds in German and English". *German Studies towards the Millennium*. Ed. Christopher Hall & David Rock. Bern etc.: Lang. 241–260.

Hohenhaus, Peter. 2005. "Lexicalization and institutionalization". In: Štekauer & Lieber 2005. 353–373.

Hohenhaus, Peter. 2006. "*Bouncebackability.* A web-as-corpus-based study of a new formation, its interpretation, generalization/spread and subsequent decline". *SKASE Journal of Theoretical Linguistics* 3: 17–27.

Jarema, Gonia. 2006. "Compound representation and processing: A cross-linguistic perspective". In: Libben & Jarema 2006. 45–70.

Jarema, Gonia & Gary Libben. eds. 2007a. *The Mental Lexicon: Core Perspectives*. Amsterdam: Elsevier.

Jarema, Gonia & Gary Libben. 2007b. "Introduction: Matters of definition and core perspectives". In: Jarema & Libben 2007a. 1–6.

Jescheniak, Jörg. 2002. *Sprachproduktion. Der Zugriff auf das lexikale Gedächtnis beim Sprechen.* Göttingen: Hogrefe.

Kastovsky, Dieter. 1978. "Zum gegenwärtigen Stand der Wortbildungslehre des Englischen". *Linguistik und Didaktik* 36: 351–366.

Kastovsky, Dieter. 1982. *Wortbildung und Semantik*, Düsseldorf/Bern/München: Schwann-Bagel & Francke.

Kastovsky, Dieter. 1986. "The problem of productivity in word-formation". *Linguistics* 24: 585–600.

Knobel, Mark, Matthew Finkbeiner & Alfonso Caramazza. forthcoming. "The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production". *Cognitive Neuropsychology.*

Krug, Manfred. 2003. "Frequency as a derminant in grammatical variation and change". *Determinants of Grammatical Variation in English*. Ed. Günter Rohdenburg & Britta Mondorf. Berlin etc.: Mouton de Gruyter. 7–67.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites.* Stanford, CA: Stanford UP.

Leech, Geoffrey. 1981. *Semantics*. 2nd ed. Harmondsworth: Penguin.
Lehrer, Adrienne. 1996. "Identifying and interpreting blends: An experimental approach". *Cognitive Linguistics* 7: 359–390.
Leisi, Ernst. 1975. *Der Wortinhalt: Seine Struktur im Deutschen und Englischen*. 5. Aufl. Heidelberg: Winter.
Leisi, Ernst. 1985. *Praxis der englischen Semantik*. 2. Aufl. Heidelberg: Winter.
Levelt, Willem J.M. 1989. *Speaking: From Intention to Articulation.* Cambridge, MA/London: MIT P.
Libben, Gary. 1994. "How is morphological processing achieved?". *Language and Cognitive Processes* 9: 369–391.
Libben, Gary. 2006. "Why study compound processing? An overview of the issues". In: Libben & Jarema 2006. 1–22.
Libben, Gary, Bruce L. Derwing & Roberta G. de Almeida. 1999. "Ambiguous novel compounds and models of morphological parsing". *Brain and Language* 68: 378–386.
Libben, Gary, Martha Gibson, Yeo Bom Yoon & Dominiek Sandra. 2003. "Compound fracture: The role of semantic transparency and morphological headedness". *Brain and Language* 84: 50–64.
Libben, Gary & Gonia Jarema, eds. 2006. *The Representation and Processing of Compound Words*. Oxford: Oxford UP.
Lipka, Leonhard. 1977. "Lexikalisierung, Idiomatisierung und Hypostasierung als Probleme einer synchronen Wortbildungslehre". *Perspektiven der Wortbildungsforschung. Beiträge zum Wuppertaler Wortbildungskolloquium vom 9.–10. Juli 1976*. Ed. Dieter Kastovsky & Herbert E. Brekle. Bonn: Bouvier. 155–164.
Lipka, Leonhard. 1981. "Zur Lexikalisierung im Deutschen und Englischen". *Wortbildung*. Ed. Leonhard Lipka & Hartmut Günther. Darmstadt: Wissenschaftliche Buchgesellschaft. 119–132.
Lipka, Leonhard. 1987. "Word-formation and text in English and German". *Neuere Forschung zur Wortbildung und Historiographie der Linguistik: Festgabe für Herbert E. Brekle zum 50. Geburtstag*. Ed. Brigitte Asbach-Schnitker & Johannes Roggenhofer. Tübingen: Narr. 59–67.
Lipka, Leonhard. 1992. "Lexicalization and institutionalization in English and German or: *Piefke*, *Wendehals*, *smog*, *perestroika*, *AIDS* etc.". *Zeitschrift für Anglistik und Amerikanistik* 40: 101–111.
Lipka, Leonhard. 1999. "*Blairites, Teletubbies, Spice Girls* and *wheelie bins* – Neologisms, the word of the year, and the nomination-function of 'words'". *Form, Function and Variation in English: Studies in Honour of Klaus Hansen.* Ed. Uwe Carls & Peter Lucko. Frankfurt a.M.: Lang. 41–48.
Lipka, Leonhard. 2000. "English (and general) word-formation – The state of the art in 1999". *Anglistentag 1999 Mainz: Proceedings.* Ed. Bernhard Reitz & Sigrid Rieuwerts. Trier: WVT. 5–20.
Lipka, Leonhard. 2002. *English Lexicology: Lexical Structure, Word Semantics and Word-formation.* Tübingen: Narr.
Lipka, Leonhard, Susanne Handl & Wolfgang Falkner. 2004. "Lexicalization & institutionalization. The state of the art in 2004". *SKASE Journal of Theoretical Linguistics* 1: 2–19.

Locke, John. [1698] 1975. *An Essay Concerning Human Understanding*. Ed. Peter Nidditch. Oxford: Clarendon.
Marchand, Hans. 1969. *The Categories and Types of Present-day English Word-formation: A Synchronic-diachronic Approach*. 2nd ed. München: Beck.
Maxwell, Kerry. 2006. *From* Al Desko *to* Zorbing: *New Words for the 21st Century*. London: Macmillan.
OED online = Simpson, John A. & Edmund S.C. Weiner, ed. 2000– (3rd ed., in progress). *The Oxford English Dictionary*. Oxford: Oxford UP. <http://oed.com/>.
Pearsall, Judy, ed. 1998. *The New Oxford Dictionary of English*. Oxford: Oxford UP.
Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Berlin etc.: Mouton de Gruyter.
Plag, Ingo. 2006. "Productivity". *The Handbook of English Linguistics*. Ed. Bas Aarts & April McMahon. Oxford: Blackwell. 537–556.
Rainer, Franz. 2000. "Produktivitätsbeschränkungen". *Morphology: An International Handbook on Inflection and Word-Formation*. Vol. I. Ed. Geert Booij, Christian Lehmann & Joachim Mugdan. Berlin/New York: de Gruyter. 877–885.
Rainer, Franz. 2005. "Constraints on productivity". In: Štekauer & Lieber 2005. 335–352.
Ryder, Mary Ellen. 1994. *Ordered Chaos: The Interpretation of English Noun-noun Compounds*. Berkeley: U of California P.
Sandra, Dominiek. 1990. "On the representation and processing of compound words: Automatic access to constituent morphemes does not occur". *Quarterly Journal of Experimental Psychology* 42: 529–567.
Sandra, Dominiek. 1994. *Morphology in the Reader's Mental Lexicon*. Frankfurt a.M.: Lang.
Sauer, Hans. 1992. *Nominalkomposita im Frühmittelenglischen: Mit Ausblicken auf die Geschichte der englischen Nominalkomposition*. Tübingen: Niemeyer.
de Saussure, Ferdinand. [1916] 1983. *Course in General Linguistics*. Ed. Charles Bally, trans. Roy Harris. London: Duckworth.
Schmid, Hans-Jörg. 2005. *Englische Morphologie und Wortbildung: Eine Einführung*. Berlin: Schmidt.
Schmid, Hans-Jörg. 2007. "Entrenchment, salience and basic levels". *The Oxford Handbook of Cognitive Linguistics*. Ed. Dirk Geeraerts & Hubert Cuyckens. Oxford: Oxford UP. 117–138.
Schmid, Hans-Jörg. forthcoming. "Conceptual blending, relevance and word-formation". *Windows to the Mind: Metaphor, Metonymy and Conceptual Blending*. Ed. Sandra Handl & Hans-Jörg Schmid. Berlin etc.: Mouton de Gruyter.
Schreuder, Robert & R. Harald Baayen. 1997. "How complex simplex words can be". *Journal of Memory and Language* 37: 118–139.
Semenza, Carlo & Sara Mondini. 2006. "The neuropsychology of compound words". In: Libben & Jarema 2006. 71–95.
Smyk-Bhattacharjee, Dorota. 2006. "Acceptance as an integral factor in the interpretation of novel words". *SKASE Journal of Theoretical Linguistics* 3: 28–36.

Štekauer, Pavol. 1998. *An Onomasiological Theory of English Word Formation*. Amsterdam: Benjamins.
Štekauer, Pavol. 2005. *Meaning Predictability in Word Formation: Novel, Context-free Naming Units*. Amsterdam: Benjamins.
Štekauer, Pavol & Rochelle Lieber, eds. 2005. *Handbook of Word-formation*. Dordrecht: Springer.
Sykes, J.B., ed. 1976. *The Concise Oxford Dictionary of Current English*. 6th ed. Oxford: Oxford UP.
Taft, Marcus & Kenneth Forster. 1976. "Lexical storage and retrieval of polymorphemic and polysyllabic words". *Journal of Verbal Learning and Verbal Behavior* 15: 607–620.
Thompson, Della. 1995. *The Concise Oxford Dictionary of Current English*. 9th ed. Oxford: Oxford UP.
Tuggy, David. 2005. "Cognitive approach to word-formation". In: Štekauer & Lieber 2005. 233–265.
Ungerer, Friedrich. 2002. "The conceptual function of derivational word-formation in English". *Anglia* 120: 534–567.
Ungerer, Friedrich. 2007. "Derivational morphology and word-formation". *The Oxford Handbook of Cognitive Linguistics*. Ed. Dirk Geeraerts & Hubert Cuyckens. Oxford: Oxford UP. 991–1025.
Ungerer, Friedrich & Hans-Jörg Schmid. 1998. "Englische Komposita und Kategorisierung". *Rostocker Beiträge zur Sprachwissenschaft 5: Kognitive Lexikologie und Syntax*. Ed. Friedrich Ungerer. Rostock: Universität Rostock, Philosophische Fakultät. 77–98.
Ungerer, Friedrich & Hans-Jörg Schmid. 2006. *An Introduction to Cognitive Linguistics*. 2nd ed. London etc.: Pearson Longman.
Vygotsky, Lew Semjonowitsch. 1962. *Thought and Language*. Cambridge, MA: MIT P.
Warren, Beatrice. 1978. *Semantic Patterns of Noun-noun Compounds*. PhD Diss. University of Göteborg.
Zimmer, Karl E. 1971. "Some general observations about nominal compounds". *Working Papers on Language Universals* 8: 3–20.
Zwitserlood, Pienie. 1994. "The role of semantic transparency in the processing and representation of Dutch compounds". *Language and Cognitive Processes* 9: 341–368.

München Hans-Jörg Schmid