SOUND SYMBOLISM IN A USAGE-DRIVEN MODEL

ANATOL STEFANOWITSCH

Abstract

Sound-symbolism is often assumed to be a universal, extra-linguistic property of the human mind. Sound-symbolic linguistic units (variously referred to as phonesthemes, submorphemes, or word-affinities) are seen as a consequence of this extra-linguistic property. Under this assumption, the existence of sound-symbolic units challenges the arbitrariness of the linguistic sign. The present paper argues that there is no reason to assume that sound symbolism is extra-linguistic. Instead, it can be seen as an expected outcome of schema abstraction over language-specific vocabulary structure. Consequently, most types of sound-symbolic units do not challenge the arbitrariness of the sign.

1. Introduction

All current morphological theories share two fundamental assumptions: First, that the morpheme is the smallest meaningful unit of linguistic structure, and second, that the link between form and meaning is arbitrary (at least at and below the word level). Both assumptions have served well as axioms on which insightful descriptions of linguistic structure and linguistic meaning can be built.

However, there is a vast literature on sound symbolism, which purports to challenge at least the first of these two premises, and often also the second. Sound symbolism here refers to recurrent form-meaning correlations below the level of what is traditionally considered to be a morpheme. Textbook examples from English are *gl-* which is claimed to be associated with the domain of LIGHT or VISION, as in *glitter, glow, gleam, glance, glare*, etc., and *fl-* which is said to be associated with RAPID MOVEMENT, as in *flick, flicker, flit, flip*, etc. I will refer to such correlations as *sound-symbolic structures*.

Sound-symbolic structures differ from morphemes in several ways, which I will discuss below. Most importantly, they are accompanied by what I will refer to as the *sound-symbolic sensation*, by which I mean the feeling that the form of a particular word somehow 'fits' its meaning.

Previous approaches often take the sound-symbolic sensation as a given, and treat sound-symbolic structures as a result of this sensation. They assume that sound-

symbolic sensations have an extra-linguistic basis, that they are a fundamental (and universal) property of the human mind. In this view, the sound-symbolic sensation directly shapes the lexicon of a language, resulting in the existence of sound-symbolic structure.

This paper takes the opposite view, arguing that sound-symbolic sensations derive from the sound-symbolic structures found in a given language. Taking a usage-driven approach to language (Langacker 1987, 1990, 1991, cf. also Bybee 1988, Lamb 1999), I will show that sound-symbolic sensations are an expected consequence of language use, driven by the same linguistic processes that underlie phonology and morphology.

This perspective raises the question where sound-symbolic structures come from in the first place. Although this is not the main focus here, I will offer some suggestions, arguing that they are typically the result of diachronic processes, such as divergence from a common root or the blending of existing lexical items.

2.   Sound symbolism: the phenomenon

The term *sound symbolism* is generally used as a cover term for both *onomatopoeia* (sometimes also referred to a *primary sound symbolism*) and *phonesthesia* (sometimes referred to as *secondary sound symbolism*).

Defining the first of these is relatively unproblematic: it is the imitation of an environmental sound using linguistic means, or, more specifically, the conventionalized representation of such a noise within the phonological system of a given language (cf. Hinton et al. 1994: 3, cf. also Rhodes 1994 for a discussion of the highly structured nature of sound symbolism). Onomatopoeia does not play an important role in this paper (but cf. Langacker 1987: 79f. for a treatment within Cognitive Grammar).

The term *phonesthesia* is much more difficult to define. Bolinger (1981) defines it as follows: 'words clustering in [semantic] groups with a vague resemblance in sounds' (Bolinger 1981: 130), or 'the grouping of similar meanings around similar sounds', distinct from morphological units in that they are 'too hazy to carve out as a definite morpheme' (Bolinger 1981: 130), and, more importantly in that they trigger 'a vague feeling of "aptness" within a given speech community' (Reay 1994: 4064), i.e. that 'the outer and inner affinity between such vocables is intuitively felt by the ordinary members of the speech community' (Jakobson and Waugh 1987: 201).

Thus, the term phonesthesia is essentially applied to two different phenomena. One is a fact about the lexicon of a given language; for example, it is a fact about English that many verbs of light emission and reflection begin with the sequence [gl]. This is what I refer to as *sound symbolic structure*. The other is a fact about a particular speaker's conscious experience of language; for example, a particular speaker may feel that the sequence [gl] 'fits well' with the meaning LIGHT. This is what I refer to as the *sound symbolic sensation* or *association*, regardless of whether it is limited to an individual or shared across many members of a speech community.

Any theory of sound symbolism must therefore answer the question of where the sound symbolic feeling and sound symbolic structures come from. Since the two are logically independent of each other, a theory of sound symbolism must also answer the question of how the two are related. This paper cannot review in full the vast literature

on sound-symbolism. Instead, I will focus on the general answers that have been suggested for both these questions (for a overview of the history of ideas about sound symbolism, see Jakobson and Waugh 1987, ch. 4).

2.1. Sound symbolic structure

Many claims about the origin and the nature of sound-symbolic structure is based on a particular example of vowel symbolism that is generally referred to as *size-sound symbolism*, i.e. cases where some aspect of vowel quality (typically height, and sometimes frontness) is said to correlate with the size of a given referent. The first part of this section will describe this phenomenon in some detail. The second part will then introduce in more detail the type of consonant symbolism exemplified above in connection with LIGHT and MOVEMENT.

The first mention of size-sound symbolism is probably in Plato's Cratylos (§426c), where Socrates says of a hypothetical name-giver that he used α for the "expression of size" because it is a "great letter." It was introduced into modern linguistics by Jespersen (1922 [1967]). The basic claim is that high front vowels like English [i] "often indicate what is small, slight, insignificant or weak" (ibid.: 283), while [u], [o], and [a] are associated with "bigger things" (ibid.: 284, cf. Fónagy 1963: 12-21 for the history of this idea).

The correlation between size and vowel height and frontness seems to be supported by a vast experimental literature on association tasks which shows that in many unrelated languages speakers associate smallness with high front vowels and bigness with low/back vowels (Chastaing 1965 for French, Newman 1933 and Sapir 1929 for (American) English, Tarte 1974 for Czech, Fónagy 1963 for Hungarian, Fischer-Jørgensen 1967 for Danish).

However, it is less clear whether this associative link (which can be experimentally induced using nonce words or isolated vowels) actually plays a role in language structure. On the one hand, it seems that there may be correlates of this associative link *within* a given language. Jespersen (1922) lists several pages of words in which high front vowels correlate with small size from various (mostly Indo-European) languages, drawing particular attention to the fact that diminuitive suffixes in these languages all contain high front vowels. On the other hand, attempts to establish a link between the sound-symbolic sensation and sound-symbolic structure cross-linguistically have so far failed; systematic typological studies show no correlation between high/front vowels and diminuitives and low/back vowels and augmentatives (cf. e.g. Ultan 1978, Bauer 1996). Moreover, even where systematic correlations between size and sound exist, they do not necessarily match the sound-symbolic sensations established in the experimental literature. Diffloth (1994) describes size-sound symbolism in Bahnar (Mon Khmer) expressives. In Bahnar, bigness is systematically associated with high vowels, and smallness with low vowels, as evidenced by a large set of contrasting pairs which differ only in vowel height, e.g. [jəhuuh] 'sth large and gaping' vs. [jəhɔɔh] 'sth. small and gaping' (Diffloth 1994: 111).

In sum, the experiments mentioned seem to provide at least tentative evidence for certain language-independent sound-symbolic sensations. However, the corresponding sound-symbolic *structures* are far from universal.

Matters are even less clear with respect to consonant symbolism. Here, most

attention has been focused on initial and final clusters. There is solid evidence for sound symbolic structure within individual languages. Take for example what is without doubt the most famous English phonestheme: [sl], which is associated with SLIDING MOVEMENT, as in *sledge, slick, slip*, etc.; WETNESS, as in *slime, slosh, sludge*, etc.; INDOLENCE, as in *slack, sleazy, slob*, etc. (cf. e.g. Käsmann 1992, cf. also Reay 1994: 4064). Other frequently mentioned examples are [gl] for VISION, mentioned in the Introduction, and [sw] for SMOOTH/WIDE-REACHING MOVEMENT, as in *sweep, swerve, swivel, swish*, etc. (cf. Reid 1967: 39). As a final example, take [st] which is associated with either MOVEMENT, as in *step, stomp, stray, stride, stroll*, etc., or with the opposite concept of ARREST, as in *stall, stay, stop*, etc. (cf. Bolinger 1965: 223). There is some experimental evidence that speakers are aware of such sound-symbolic structures in their language and are able to draw on them when asked to create nonce words expressing some concept, or assigning a meaning to such nonce forms (cf. Abelin 1999 for a series of experiments with Swedish speakers).

However, it is even clearer in the case of such consonant clusters that such associations are limited to individual languages. Take the English phonestheme [gl], associated with LIGHT or VISION. A similar association can be found (although to a lesser degree in very closely related languages, such as German, or Swedish). However, it cannot be found even in slightly more distantly related languages, let alone cross-linguistically.

2.2. The sound symbolic sensation

It is often assumed in the relevant literature, that the sound symbolic sensation originates outside of language, i.e. that it is not itself in need of a linguistic explanation, but can be taken as given and then be used to explain the existence of sound symbolic structures.

The external motivation for the sound symbolic sensation is sought in natural connections between sound and meaning. The connection is seen to be provided either by synesthesia or by kinesthetic experience. The first type of explanation (cf. Peterfalvi 1970, Abelin 1999: 43f) assumes that sensory input from one modality can be generalized to other modalities due to inherent properties of the brain. Such synesthetic generalization are assumed to underlie the experiments discussed in the preceding section. The second type of explanation is very similar, looking for bodily experiences during speech production that may explain such generalizations. For example, [i] is said to symbolize smallness because of the small aperture between tongue and palate, while [a] symbolizes largeness due to the large aperture (such explanations grounded in articulatory sensations can already be found in Plato's *Cratylos,* see Fonágy 1963: 12-37 for the history of such explanations, cf. also Diffloth 1994: 107f).

However, there is also a line of explanation which views sound symbolism at least to some degree as a linguistic phenomenon. This explanation, proposed by Bolinger in various publications, assumes that speakers approach their language with the assumption that "if things sound alike, they must mean alike, and if they mean alike, they ought to sound alike" (Bolinger 1994: 29, cf. also 1981: 131).

In a series of publications, Linda Waugh has laid out an approach to the lexicon which is based on this assumption, which she calls the *principle of isomorphism*, and which she states as follows:

[S]ameness in form from one sign to another signals sameness of meaning and difference of form signals difference of meaning (Waugh 1994: 56)

Waugh's approach accommodates sound symbolism within a general theory of form-meaning relations: in her opinion, the principle of isomorphism is the basis for any analysis of words into smaller units, most importantly, into those units which are traditionally called *morphemes* in a

> set of words like *water, watery, waterfall, rain, rainy, raindrop, snow, snowy, snowshoe*, whereby, for example, *water, watery,* and *waterfall* share a meaning, which is formally represented by the same form *water* in all three words. And likewise with the *-y* of *watery, rainy,* and *snowy*. *Watery* is diagrammatically related to *waterfall* on the one hand because of the common root *water* and to *rainy* on the other hand because of the common derivational suffix *-y*; *watery* is motivated relative to these and other words through the recurrence of the same form–meaning relationship (ibid.: 57).

This analysis is interesting because it takes a perspective on morphology which is radically different than that which we find in current morphological theory. In a sense, it is an extension of the structuralist principle that 'every word gets its meaning through its relations with any other word in the system': every word gets its internal structure through its relationship to every other word in the system.

Waugh argues that the isomorphic principle can also account for those aspects of internal word structure which cannot be accounted for by traditional morphology, most importantly phonesthemes (ibid.: 59). Thus, it is implied in her account that phonesthemes are not qualitatively different from morphemes. Waugh then goes on to make several interesting points, which are usually neglected in discussions of phonesthesia: since they are no different from morphemes, they are subject to the same constraints. For example, they will display various degrees of semantic coherence (ranging from a unitary meaning over polysemy, to homonymy), and various degrees of compositionality.

Note that Bolinger's and Waugh's line of explanation allows for sound symbolism as a shaping force in the lexicon of a language, since once a cluster of word-affinities is established, it may become moderately productive in the sense that it may exert pressure on existing forms to conform to the pattern (Waugh 1994: 61). However, the question remains, how clusters of related words enter the vocabulary of a language in the first place. Note also that ultimately, their explanation is still extra-linguistic: their principle of isomorphism is assumed to be a basic fact of human cognition.

The next section will introduce the Cognitive Linguistic conception of language a usage-driven network. Starting out from this conception, the subsequent section will then develop a theory of the sound-symbolic sensation which shares certain aspects with Bolinger's and Waugh's theory, but which does not assume the existence of an isomorphic principle. Instead, the sound-symbolic sensation, and with it the ability to draw on sound-symbolic associations in producing or interpreting nonce-forms follow directly from the workings of the linguistic system.

3. A usage-driven network model of language

3.1. The linguistic system as a network

Cognitive approaches to language typically assume the conception of the linguistic system as a dynamic network whose development and structure are driven by usage (Langacker 1987, 1990, Bybee 1988, Lamb 1999). These models share certain fundamental properties that distinguish them radically from the more traditional item-and-process or item-and-arrangement models as well as current constraint-based theories. I will mainly use the version of the model developed by Langacker, but it is the fundamental properties shared by all models rather than the specific details of any one of them that will provide the theoretical apparatus for this investigation. Since, as I will argue, sound symbolism shares properties of both phonology and morphology, I will concentrate on these two areas after giving a general overview of cognitive linguistics.

3.1.1. Linguistic units in a network model. As already mentioned, cognitive linguistics assumes that the linguistic system has the form of a network, i.e. that it consists of nodes and connections between these nodes. In this assumption, cognitive linguistics is in full agreement with current psycholinguistic and AI-based theories of language. There seems to be agreement that such a network can ultimately be correlated with the human central nervous system, or more precisely, the brain (Lamb 1999 is very explicit on this point, and has the most detailed proposal of what the exact correspondences are, but see also Langacker 1990: 10, 1991: 525).

Crucially, such a network does not contain linguistic units in the sense of formal entities that can be operated on by rules or constraints. Instead, linguistic units are seen as sets of connected nodes.

Morphemes, lexemes and syntactic constructions are seen as bipolar connections between a form and a meaning. The form is referred to as the *phonological pole*. In the case of spoken language it ultimately corresponds to an articulatory or auditory routines, cf. Bybee 1994, Lamb 1999: 126ff). The meaning is referred to as the *semantic pole*. It has the form of an open-ended set of connections within a knowledge network (cf. Langacker 1987: 154ff) which is ultimately tied to embodied experience, e.g. in the form of sensory information (cf. Lakoff 1987 and Lamb 1999, Ch. 9). For example, the word 'dog' is represented in the linguistic system as a node which has connections both to the articulatory/auditory routines that produce/recognize the sound shape [dçg] and to our knowledge about dogs, including the sensory information of what a dog looks like, feels like, sounds like, what it typically does, etc. Bipolar units are referred to as *symbolic units*. There are various ways of diagramming them. All of these have advantages and drawbacks. I follow Langacker's notation, mainly because it is the most compact, but I will occasionally make reference to other notations. In the notation I will use, the phonological pole is shown orthographically in lower case (or in IPA where necessary) followed by a slash, followed by the meaning in upper case (cf. Langacker 1987: 76f.). The meaning will be given with various degrees of detail and completeness, depending on the relevance. For example, *dog* could be represented as [dɔg/ANIMAL KEPT BY HUMANS FOR COMPANIONSHIP, HUNTING, GUARDING, ...] or simply as [dog/DOG].

In addition to nodes integrating phonology and semantics, the network may also

contain nodes that integrate only phonological information, or only semantic information. (Bybee 1988: 125f., Langacker 1990: 4). Purely semantic units play no role in this paper, but various types of phonological units will be discussed in detail in the next section.

Note that the account sketched out so far makes some explicit claims about storage. A linguistic unit is stored in the network in terms of a node integrating information from different parts of the network. Lamb (1999) points out that this means there is both a local and a distributed representation, depending on the perspective we take. The information is distributed throughout the network, but the node integrating all the connections to this information is of course local.

Finally, units in the network have different degrees of entrenchment. This is a consequence of frequent activation: the more often a certain node is activated the more easily it will be activated in the future (cf. Langacker 1987: 59). For a purely phonological unit the degree of entrenchment is dependent only on the frequency with which that unit is activated. For a symbolic unit, a high degree of entrenchment may be due either to the high token frequency of the phonological pole or to the high familiarity of the semantic pole (i.e. the concept to which it corresponds).

3.1.2. Schemas and complex units. Finally, the network contains units of different degrees of complexity and different degrees of abstractness. This is due to the fact that a set of linguistic units which are frequently activated together achieves unit status itself.

Complex units emerge where two or more simple units are frequently activated in sequence. For example, the units [dog/DOG] and [biscuit/BISCUIT] are often activated together. Even though the sequence of the two simple units, *dog biscuit*, is relatively transparent both formally and semantically, frequent co-activation will cause this sequence itself to become a unit in the network, which can be represented as [dog biscuit/DOG BISCUIT]. Once a complex unit is established, it can, but need not, begin to diverge formally and semantically from the simple units contained in it (Lamb 1999: 163ff., cf. also Langacker 1987; 23f).

Schematic units emerge where a set of complex units shares some properties, but diverges in others. The schematic unit (or *schema*) captures the commonalties among the different complex units, abstracting away from the differences. Schematic units do not differ qualitatively from the units over which they are abstracted, but they differ with respect to granularity, i.e. with respect to how much detail they specify at the semantic and the phonological pole (Langacker 1990: 4). For example the words *hot dog*, *cheese dog*, and *chili dog* all share some aspects of form and meaning, but diverge in others. Given this situation, a schema is abstracted in the network capturing the similarities on both levels. I extend Langacker's notation as follows in order to show the abstraction process:

hót dog/FRANKFURTER IN A ROLL WITH ONIONS AND KETCHUP
chéese dog/FRANKFURTER IN A ROLL WITH CHEESE
chíli dog/FRANKFURTER IN A ROLL WITH CHILI
_____

[...] dog/SAUSAGE IN A ROLL WITH ...

Above the line the actual units are shown, below the line those aspects common to all form-meaning pairs are shown as a more abstract unit of the same kind, with ... showing the underspecified aspects. Where the underspecified aspects correspond to units, I will sometimes enclose them in vertical lines.

The way that such schemas arise is crucial in understanding the sound symbolic sensation. The next section will discuss this process in detail both for purely phonological schemas and for morphological schemas.

3.2 The usage-driven nature of the network: more about schema abstraction

3.2.1. Morphological schemas. Meaningful word-internal structure emerges when a schema is abstracted over a set of linguistic units that share similarities both at the phonological and the semantic pole (cf. Bybee 1988: 127, Langacker 1990b: 278f, Lamb 1999: 212ff.). This process was briefly shown for the [|...| dog] schema above. Let us look at the process in more detail by considering the regular English plural suffix in the words *cars*, *dogs*, and *boys*. These words will initially be learnt as unanalyzed simple units. However, they share similarities at both poles. At the phonological pole, the units are similar in that they all contain a [z], preceded by a voiced element, but they differ in what precisely this voiced element is or what precedes it. At the semantic pole, they all share the notion of MORE THAN ONE, but they differ in what it is of which there is more than one. The schema that is abstracted shares the commonalties of all the units, but differs in being semantically and phonologically underspecified where the units differ from each other:

kɑːɹz / MORE THAN ONE CAR
bɔɪz / MORE THAN ONE BOY
dɔgz / MORE THAN ONE DOG
———————————————————
[..vcd]z / MORE THAN ONE ...

Of course, there are different types of plural suffixes, depending on the nature of the segment preceding the plural suffix. Separate schemas are abstracted for each type, and an even more abstract schema is abstracted over these schemas:

[..vcd]z / MORE THAN ONE ...
[..vcl]s / MORE THAN ONE ...
[..sbl]ɪz / MORE THAN ONE ...
———————————————————
[...][alv frc] / MORE THAN ONE ...

These underspecified representations allow us to recognize a form as a plural, and they allow us to produce novel plurals.

The type of diagramming chosen here fails to make explicit some crucial aspects of schema abstraction, which can easily be overlooked or misinterpreted.

First, and most importantly, the different forms are not represented as completely separate units in the network. Instead, the similarities between different forms are due to

the fact that they share some of the same network structure. For example, there is only one auditory and one articulatory routine corresponding to [z], and the different plural forms shown above are similar in that they all have connections to these routines. Likewise, the concept of MORE THAN ONE is only represented once in the network, and the units are similar in that they all share connections to this concept (which is grounded in our experience of multiplicity). This fact of shared network structure is not directly captured by Langacker's representations (or by Bybee's; as far as I can tell, Lamb's relational network notation is the only one that directly captures it). However, all network models agree on the issue itself. We can think of the symbols on either side of the slash as representing the *links* to the corresponding network structures.

Second, this fact of shared network structure means that schemas are not represented in the network separately from the concrete units over which they are abstracted. Instead, a schema simply *is* the set of shared nodes and connections. It emerges as a unit because the shared structure will be activated more frequently than the non-shared structures. Thus the schema is immanent in its manifestations (Bybee 1988: 124, Langacker 1990a: 10, Lamb 1999: 265ff). For example, the combination of nodes [z/MORE THAN ONE] is activated every time one of the words *dogs, boys, cars,* or any other plural form of this type is activated. Its frequency of activation is the sum of the frequencies of activation of every individual plural form.

Third, connections may have different degrees of entrenchment, and therefore schemas abstracted over such connections can also have different degrees of entrenchment. I will return to this point in the discussion of sound symbolism below. Different degrees of entrenchment have consequences for example for the issue of productivity (excellent discussions of this can be found in Bybee and Moder 1983; Bybee's notation is well suited to showing different degrees of entrenchment).

Finally, such schemas will typically be abstracted during language acquisition. However, cognitive linguistics assumes that the linguistic system is continually being shaped by usage throughout the life of a speaker (i.e. it never reaches a 'steady state'), thus schemas can be abstracted at any point. For example, schema abstraction is also the process by which loan morphology enters a language (see Stefanowitsch and Rohde 1999, in preparation).

3.1.2. Phonological schemas. The precise nature of the minimal phonological unit (something like the phoneme/allophone) is not relevant in the context of this paper, thus I will not discuss it in any detail. Essentially, in cognitive phonology it is assumed that linguistic units are initially represented holistically, i.e. that they are nodes connected to a fully specified auditory/articulatory template on the one hand and to an unanalyzed portion of semantic knowledge on the other (Bybee 1994: 292f, cf. also Langacker 1987: 328ff, 1990: 272ff). Frequently co-occurring aspects of such auditory/articulatory routines can then achieve unit status; thus portions of varying size of such routines can achieve independent unit.

Complex phonological units are not essentially different from complex morphological units. What happens is that connections are established between phonologically similar items that do not share semantics, e.g. *practice, present, pretty, private,* and *problem*. From these, a more abstract schema [pɹ.../Ø] can be abstracted, capturing the commonalties at both poles. Since there are no similarities at the semantic

pole, the schema will contain no semantics:

pɹæktɪs / PRACTICE
pɹɛsnt / PRESENT
pɹɪɾi / PRETTY
pɹaɪvɪt / PRIVATE
pɹɑ:bləm / PROBLEM
————————————

pɹ... / Ø

Depending on the nature of the units over which they were abstracted, such units may be of all sizes and degrees of abstractness. Apart from the fact that they have no entrenched set of connections to a semantic pole, they differ from typical morphological schemas in that they do not necessarily result in the full analyzability of a unit in terms of smaller units. Thus, [pɹ] may be a unit in the examples above, but [ɛsnt] or [ɪɾi] are not.

4. Sound symbolism: the analysis

4.1 The sound-symbolic sensation in a usage-driven network
    The process whereby sound-symbolic associations are formed is not qualitatively different from the processes that lead to the abstraction of phonological or morphological schemas. The sound-symbolic association is first and foremost the result of schematic units which emerge under the condition that there is a set of units with partial identity at both the semantic and the phonological pole. For example, as already mentioned the words *slick*, *slide*, and *slip* all share the initial phoneme sequence [sl] and they all share some aspect of SMOOTHNESS. Thus, a schema will be abstracted as shown (in the by now familiar notation):

slick / MAKE STH. SMOOTH AND SHINY
slide / MOVE OVER SMOOTH SURFACE, CONTINUOUS CONTACT
slip / MOVE QUICKLY, UNEXPECTEDLY OVER SMOOTH SURFACE
————————————————————————————————————

sl ... / SMOOTH ...

This schema is itself only a small part of the larger network of associations mentioned in Section 2.1.
    Thus, sound-symbolic associations are an automatic and fully expected by-product of schema abstraction. However, the analysis here leaves two questions unanswered: first, why do we treat sound symbolism as different from morphology, given that they seem to have so much in common, and second, given that sound-symbolic associations are an outcome rather than a source of sound-symbolic structures, where do these structures come from in the first place.

4.2 Similarities and differences between sound symbolism and morphology

Given that the process through which phonesthemes emerge is so similar to the process through which canonical morphemes emerge, what, if any, are the differences between morphology and sound-symbolic structure? The differences or similarities could be found in three areas: (i) systematicity and strength of the form-meaning correlation (i.e. degree of entrenchment of a schema); (ii) degree of analyzability; and (iii) the sound-symbolic sensation.

The first issue to be dealt with is, how systematic the kinds of form-meaning recurrences are, that are typically given as examples of phonesthemes. Collecting a number of words that share some aspect of form and meaning does not prove that a hypothetical schema capturing the recurring form-meaning correlation will be entrenched enough to have an effect in the linguistic system, or more precisely, to cause a sound-symbolic association.

The typical method of establishing an argument for a sound-symbolic structure is to pick out a set of words that share some formal feature and construct a semantic link between them (Dwight Bolinger was a master of this method). The problem with this method is that it often shows nothing more than the ingenuity of the linguist in finding such sets of words. For example, I could argue that there is a phonestheme [pɹ] in English, meaning NEAT OR PLEASANT APPEARANCE, and cite the words *preened*, *preppy*, *pretty*, *prim* and *proper*, *pristine*, and *primped* as evidence. Alternatively, I could argue that [pɹ] is associated with the meaning AMORAL, and cite *promiscuous*, *prostitute*, *prurient profligate*, and *depraved*. Finally, I could argue that [pɹ] is associated with RELIGION, and cite *praise*, *pray*, *preach*, *prebendary*, *priest*, *presbytery*, *prior*, *profess*, *profanity*, *Promised Land*, *prophet*, *proselyte*, and even *Pre-Raphaelite*. An average-size dictionary like the LDCE has 751 entries with the initial cluster [pɹ], more than enough to contain some sets of semantically related words, especially where the semantic relation is as loose as in the cases just cited.

In order to distinguish such cases from linguistically significant cases of sound-symbolism, the strength of the correlation must be determined. Given the usage-based nature of the linguistic system, this strength will depend on the frequency with which a combination of sound and meaning is activated. We can assume that the entrenchment of a schema can be measured in terms of *cue validity*. A schema will be strongly entrenched if the two poles are co-activated frequently, and if each pole is co-activated with other poles infrequently.

As an example, consider the phonestheme *gl-*, one of the textbook examples mentioned in the Introduction. Fromkin & Rodman (1974: 3) note that "[i]n English, there are many words beginning with *gl-*, such as *glare, glint, gleam, glitter, glossy, glance, glimmer, glimpse, glisten*, all of which have to do with sight." Note that Fromkin & Rodman's semantic characterization is very vague. A look at their list shows that it actually consist of two different groups of words: verbs which denote EMISSION OF LIGHT (e.g. *glitter*) and words which denote WAYS OF LOOKING (e.g. *glare, glance*). I will deal with the first set here.

The first step consists in calculating the cue validity, i.e. on the one hand the degree to which the activation of the initial cluster [gl] predicts the activation of the semantic feature 'emitting light', and on the other hand the degree to which the activation of the semantic feature 'emitting light' predicts the activation of word initial [gl]. In order to

do so, we have to calculate (i) how many words an average speaker of English knows which have LIGHT EMISSION as one of their semantic features, (ii) how many words an average speaker of English knows which start with *gl-*, and (iii) how many words are there which combine both properties.

In order to answer the first two questions, we have to make an assumption about the words which an English speaker is likely to know, and about the words which are likely to be entrenched enough to participate in schema abstraction. I will assume that the frequency list in Kucera & Francis (1967) is a good indicator of this, i.e. I will only include words in the following discussion which appear at least once in this list.

A cross-check of entries in the ODDT yields a list of 21 verbs which seem to involve the meaning LIGHT EMISSION. This list is shown here (the figures are the corresponding frequencies from Kucera & Francis (1967):

blind	(56)	blaze	(15)	radiate	(7)
flash	(52)	gleam	(14)	twinkle	(5)
beam	(40)	glisten	(14)	flicker	(4)
shine	(35)	dazzle	(13)	glimmer	(4)
glow	(33)	glitter	(12)	shimmer	(4)
glare	(20)	sparkle	(10)	scintillate	(1)
flare	(17)	glint	(9)	coruscate	(0)

On the basis of this list, we can answer the first question. An average speaker of English knows about 20 light emission verbs (although, of course, matters are more complex that that: since CG does not accept the notion of a clear dividing line between syntax and the lexicon, there are multi-word expressions which would have to be considered here as well (e.g. *emit light*, *give off light*). However, no frequency data is available for such collocations, so they will be ignored here). As to the second question, Kucera & Francis (1967) lists 43 lexemes beginning with *gl-* (excluding proper names). The third question can also be answered on the basis of the list above: there are seven light emission verbs that begin with [gl].

We can now answer the question of cue validity. Going by type frequency, we see that 7 out of 21 light emission verbs begin with *gl-*, thus the likelihood of activation (the cue validity) is 33.3%. Going by token frequency, the percentage is roughly the same: 29.0% (106 out of 365). These numbers show that the form-meaning link is far from absolute, but it is certainly above chance. In order to answer the second question, we have to count how many other words there are which begin with *gl-*. Excluding proper names, Kucera & Francis contains 43 (lexical) words of this kind, which have an overall token frequency of 662. Going by type frequency, we can see that the likelihood that *gl-* activates the meaning 'emission of light' is 7 out of 43, or 16.3%. Token frequency gives us about the same figure: 106 out of 662, or 16.0%.

How does this compare to some clear cases of morphology? First take the example of [ʌn] as in *untie, unimaginative, unhappy*. I am assuming that the most abstract schema for this form is [un-/NEGATIVE OF]. Of 758 word forms listed in Kucera and Francis beginning with [ʌn], 653 are manifestations of the unit [un-/NEGATIVE OF] (the exceptions mainly being words beginning with *under-*). Thus, the form *un-* predicts the meaning NEGATIVE OF with a cue validity of 86.15%. Looking at the opposite direction,

there are 818 word forms evoking the meaning OPPOSITE, 653 of which contain the form *un-*. Thus, the meaning NEGATIVE OF predicts the form *un-* with a cue validity of 79.83%.

Thus, there are clear quantitative differences between a fully productive morpheme like *un-* and a phonestheme like *sl-*. However, these differences alone cannot serve to distinguish sound symbolism from morphology: other morphemes look much less clear in terms of cue validity. For example, the initial sequence [baɪ] clearly has morphemic status, corresponding to the schemas [baɪ / TWO, DOUBLE], as in *bifocals, bisect, bicycle,* etc., and [baɪ / LESS IMPORTANT] as in *by-product, by-election, bystander*, etc. However, there are 80 words beginning with this sequence in Francis and Kucera, with a total token frequency of 366, and only 38.7% of the types (or a mere 13.4% of all tokens) are words instantiating one of these schemas. Non-productive morphemes like *-age* (as in *shrinkage, breakage, footage*) or morphological remnants like *-ency* as in *tendency, complacency*, etc. would presumably fare even worse.

A clearer difference between morphological schemas and sound symbolic schemas emerges in terms of full segmentability. It has been noted that words which contain sound-symbolic structure are not typically fully segmentable into phonesthemes. In other words, sound-symbolic structure differs from morphological structure in terms of what we are left with once we remove the supposedly meaningful unit. In the case of a typical productive morpheme, the rest of the word is typically also be a recognizable morpheme (*unhappy—happy, untie—tie*, etc.). In the case of a phonestheme, the rest of the word is not recognizable as anything (*glow—*ow, gleam—*eam*, etc.). In terms of the network model this can be captured by the observation that in the case of sound symbolic structures some phoneme or cluster of phonemes may achieve unit status together with some semantic structure, so that the larger units in which they occur become essentially complex, consisting of a smaller unit and a rest that does not itself have unit status. In the case of morphology it is also the case that some phoneme or cluster of phonemes achieves unit status together with some semantic structure, but the resulting complex unit consists of several such units with no non-units.

Finally, and perhaps most surprisingly in terms of the network analysis presented here, sound-symbolic structures differ from morphological structures in that the latter do not trigger a sound-symbolic sensation. There are two potential explanations for this in the network model presented here: first, the nature of the semantic structures typically associated with phonesthemes vs. morphemes, and second, the much stronger entrenchment of morphological schemas.

The semantic concepts mentioned in discussions of sound symbolism come from a single, very basic semantic domain: bodily experience. This experience may take the form of sensory impressions (e.g. [LIGHT], or [WET]), kinesthetic impressions (e.g. [FAST MOVEMENT]), or emotional impressions (e.g. [UNPLEASANT]). All these domains share the property of being located at the periphery of the semantic network, where it is linked to the senses. They serve as input, as the concrete embodied experiences on whose basis more abstract conceptual structure is created. Since this type of semantic structure is closer to the sensory surfaces of the linguistic network, its activation creates more real and more vivid experience that the activation of abstract concepts like [OPPOSITE OF] or [TWO]. Formal units with connections to more concrete experience produce stronger reactions than those with connections to more abstract conceptual structure.

The stronger entrenchment of canonical morphological schemas may also play a role in the lack of sound-symbolic sensations triggered by them. Stronger entrenchment corresponds to a lower activation threshold, which in turn means less cognitive effort. In other words, morphological schemas are fully automated; they will be triggered without any additional activation in the form of cognitive effort. Sound-symbolic schemas, on the other hand, are much more weakly entrenched, and will require extra activation in order to be activated. This extra activation means more cognitive effort, which plausibly means more attention. Thus, paradoxically, sound-symbolic structure my be more likely to be noticed precisely because it is less entrenched.

4.3 The origin of sound symbolic structure

This paper has focused on the way in which a sound-symbolic structure (i.e. a set of word affinities) leads to the abstraction of a schema within the linguistic system of a speaker, which accounts for the feeling of aptness which such sound-symbolic structures often cause, as well as for the fact that speakers may draw on such associations in interpreting or creating nonce words. The question remains, how such sets of word-affinities arise in the first place. My contention is that they arise through diachronic processes, more specifically, through (i) divergence from a common root, and (ii) blends of existing structures.

Divergence from a common root means that a single root may diverge into several semantically related forms, which share certain formal properties because the latter were present in the common root. Such a diversification account is quite plausible for the phonestheme [gl], discussed at several points in this paper (the following discussion is based on the etymological information in the *Oxford English Dictionary*).

It seems that the seven light emission verbs, as well as many other words from the general domain of vision with initial *gl-*, can be traced down to at most four (pre-)Teutonic roots: *glare* and *glow* come from the Teutonic root *glô-* (weak grade *gla-*), which also accounts for *glass*, *gloaming*, and *gloom*. *Gleam* and *glimmer* are derived from Teutonic *glim-*, along with *glim* and *glimpse*. *Glint* and the obsolete *glent* derive from Old Teutonic *glint-/glant-*, and *glitter* and *glisten*, along with the obsolete *glise* come from pre-Teutonic *ghlei(d)-/ghli(d)-*. The divergence from these roots is due to sound changes departing from different grades of these roots, and from borrowing into English from other Germanic languages. Through this process, a few roots (which could without doubt be further reduced by tracing them back to even earlier forms) can give rise to a set of words vaguely related in form and meaning which is large enough to allow for the abstraction of sufficiently entrenched schemas.

Blends may further contribute to such sets of words: when blends are created out of concrete (non-schematic) units, such blends inherit some of the phonology and some of the semantics from these units, and will thus add to the overall correlation between certain phonological units and certain semantic units (cf. Reay 1994, Sec. 3, cf. also Kemmer, to appear, for a general treatment of blends in a network model).

Incidentally, note that the network model does not preclude the possibility that the source for some words may actually be found in pre-linguistic synesthetic associations. As was argued above, this is plausible for vowel symbolism to some degree, given the consistent association of certain concepts with high and low vowels. However, even those cases have to be mediated through the process of schema abstraction in order to

14

become linguistically relevant. This is why some languages reflect size-sound symbolism in their vocabulary, while others do not. Even assuming that speakers invent new words or shape existing words on the basis of size-sound symbolism, they have to create (and maintain) a critical number of such words in order for a general schema to be abstracted. In some languages, such a critical number may simply never be reached. This is also why languages may have size-sound symbolism that goes counter to the general trend (such as Bahnar). In these languages, there is at some point a critical number of words denoting small things but containing low vowels. This will lead to a situation where a corresponding schema is abstracted that will override any 'natural' size-sound correlation.

Finally, the most important source for sound symbolic structures from a network perspective is convergence due to system-internal pressures. The sound-symbolic process is self-accelerating once it is set off (Yakov Malkiel has shown this in several places, e.g. Malkiel 1990). In the network model we can say that as soon as a sound-symbolic schema is abstracted for whatever reason, it may become productive in the creation of novel words or the shaping of existing words.

5. Conclusion

This paper has shown that within a usage-driven network model of language, phonology, sound-symbolism and morphology can be seen as different manifestations of the same basic process of schema abstraction. In the case of phonology, complex units and schemas are abstracted over a set of phonologically similar forms that do not share any aspects of meaning. In the case of sound symbolism, there are partial, and often tenuous, similarities both on the phonological pole and on the semantic pole. Consequently, schemas are abstracted that integrate these similarities, but abstraction process does not result in fully segmentable complex units or schemas. Finally, in the case of morphology the similarities are such that strongly entrenched, fully segmentable units may emerge.

The arbitrariness of the linguistic sign is not challenged by the existence of sound symbolism, since sound-symbolic structures typically emerge without external motivation. The contention that sound symbolism does challenge the arbitrariness of the sign is based not so much on the interpretation of linguistic evidence, but on deeper concerns about the relation between language and reality.

Attempts to establish a natural connection between language and reality/immediate experience are probably as old as language itself. They can be found in all linguistic traditions that we have records of, be they the Sanskrit grammarians, the Arab grammarians, or Western linguistics going back to the Ancient Greek (cf. again Fonágy 1963). Such attempts stem not so much from a desire to account for linguistic phenomena, but from the desire to find objective truth in the world around us, some objectively given stable core of language. In this paper I have argued (as others have before me) that no such link exists, that sound symbolism is not the result of any objective correlation between language and the world, nor of any correlation between language and some universal cognitive phenomenon. Instead, it turns out to be a purely linguistic phenomenon, based on the workings of the usage-driven network which

comprises speakers' linguistic systems. Although such an analysis destroys the hope of finding any objective or at least universally valid cognitive core of sound-meaning relations, it teaches us something about our place in the world. As humans, we will never have any direct access to reality, nor do we universally share a single way of perceiving the word. However, we share the ability to create meaning where there is none; it is a universal property of our brains, something that is beyond our control, but that comes as an automatic byproduct of perceiving and talking about the world.

References

Abelin, Åsa. 1999. *Studies in sound symbolism.* (Gothenburg Monographs in Linguistics 17). Göteborg: Göteborg University.

Bauer, Laurie. 1996. No phonetic iconicity in evaluative morphology. *Studia Linguistica* 50.189-206.

Bolinger, Dwight. 1965. *Forms of English*. Cambridge, MA.

Bolinger, Dwight. 1992. Sound symbolism. In William Bright (ed.), *International Encyclopaedia of Linguistics*. Oxford: Oxford University Press

Bybee, Joan L. 1988 Morphology as lexical organization. In Michael Hammond, Michael Noonan (eds.), *Theoretical Morphology. Approaches in Modern Linguistics*. New York: Academic Press, 119-141.

Bybee, Joan L. 1994. A view of phonology from a cognitive and functional perspective. *Cognitive Linguistics* 5.4: 285-305.

Bybee, Joan L. and Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59.251-270.

Chastaing, Maxime. 1965. Dernières recherches sur le symbolisme vocalique de la petitesse. *Revue Philisophique* 155.41-56.

Diffloth, Gérard. 1994. i: *big*, a: *small*. In Leanne Hinton, Johanna Nichols, John J. Ohala (eds.) *Sound symbolism*. Cambridge, New York, Melbourne: Cambridge University Press, 107-114.

Fischer-Jørgensen, Eli. 1978. In the universal character of phonetic symbolism with special reference to vowels. *Studia Linguistica* 32.80-90.

Fónagy, I. 1963. *Die Metaphern in der Phonetik*. Mouton: The Hague.

Fromkin, Victoria and Robert Rodman. 1974. *An Introduction to Language*. New York: Holt, Rinehart and Winston.

Hinton, Leanne, Johanna Nichols and John J. Ohala (eds.)1994. *Sound Symbolism*. Cambridge: Cambridge University Press.

Jespersen, Otto. 1922. Symbolic value of the vowel *I*. Philologica 1.1-19. (Reprinted 1962, *Selected writings*, 557-577. London: George Allen and Unwin.)

Käsmann, Hans. 1992. Das Englische Phonasthem sl-. *Anglia* 110.

Kemmer, Suzanne. To appear. Blends in a usage-based model.

Kucera, Henry and W. Nelson Francis. 1967. *Computational Analysis of Present-Day American English*. Providence, RI: Brown University Press.

Lakoff, George. 1987. Women, fire, and dangerous things. What categories reveal about the mind. Chicago and London: The University of Chicago Press.

Lamb, Sydney M. 1999. Pathways of the brain. The neurocognitive basis of language. Amsterdam and Philadelphia: Benjamins.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. Vol 1. *Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 1990. The rule controversy: A cognitive grammar perspective. *CRL Newsletter* 4.3: 4-15.

Langacker, Ronald W. 1991a. *Foundations of cognitive grammar*. Vol. 2. *Descriptive applications*. Stanford: Stanford University Press.

Langacker, Ronald W. 1991b. *Concept, Image, and Symbol. The Cognitive Basis for Grammar*. Berlin & New York: Mouton de Gruyter.

Malkiel, Yakov. 1990. Studies in secondary phonosymbolism. 43-67. In Yakov Malkiel, *Diachronic*

problems in phonosymbolism. Amsterdam and Philadelphia: Benjamins.

Newman, Stanley. 1933. Further experiments in phonetic symbolism. American Journal of Psychology 45.53-75.

Peterfalvi, J.M. 1970. Les recherches expérimentales sur le symbolisme phonetique. American Journal of Psychology 65.439-473.

Reay, I.E. 1994. Sound symbolism. In R.E. Asher and Y.M. Simpson (eds.), *The Pergammon Encyclopaedia of Language and Linguistics*. Oxford: Clarendon Press.

Sapir, Edward. 1949. A study in phonetic symbolism. Selected writings, ed. by D. Mandelbaum, 61-72. Berkeley, CA. [orig. published 1929, Journal of Experimental Psychology 12.225-239]

Stefanowitsch, Anatol, and Ada Rohde. 1999. Borrowing building blocks: loan morphology and pseudoloans in a usage-driven model. Paper delivered at the 28 LASSO Meeting.

Stefanowitsch, Anatol, and Ada Rohde. In preparation. Pseudoloans in a usage-driven model. [Note: This paper was never published.]

Tarte, R.D. 1974. Phonetic symbolism in adult native speakers of Czech. Language and Speech 17.87-94.

Ultan, R. 1978. Size-sound symbolism. Universals of human language, vol. 2, ed. by J.H. Greenberg, C.H. Ferguson, E.A. Moravcsik, 527-568. Stanford: Stanford University Press.

Waugh, Linda. 1994. Degrees of iconicity in the lexicon. *Journal of Pragmatics* 22: 55-70.