

THE PSYCHOLOGICAL REALITY OF PHONAESTHEMES

BENJAMIN K. BERGEN

University of Hawai'i at Manoa

The psychological reality of English phonaesthemes is demonstrated through a priming experiment with native speakers of American English. Phonaesthemes are well-represented sound-meaning pairings, such as English *gl-*, which occurs in numerous words with meanings relating to light and vision. In the experiment, phonaesthemes, despite being noncompositional in nature, displayed priming effects much like those that have been reported for compositional morphemes. These effects could not be explained as the result of semantic or phonological priming, either alone or in combination. The results support a view of the lexicon in which shared form and meaning across words is a key factor in their relatedness, and in which morphological composition is not required for internal word structure to play a role in language processing.*

1. INTRODUCTION. Phonaesthemes (Firth 1930) are frequently recurring sound-meaning pairings that are not clearly contrastive morphemes. An example is the English onset *gl-* (Wallis 1699, Bloomfield 1933), which, like other phonaesthemes, is relatively infrequent in English, except among words with meanings related to 'vision' and 'light'. Some of these are exemplified in 1a. Another well-documented phonaestheme (Wallis 1699, Bloomfield 1933) is the English onset *sn-*, which occurs in a large number of words relating to 'mouth' and 'nose' (1b).

- (1) a. *gl-* 'light, vision' *glimmer, glisten, glitter, gleam, glow, glint, etc.*
- b. *sn-* 'nose, mouth' *snore, snack, snout, snarl, snort, sniff, sneeze, etc.*

When viewed simply as statistically aberrant distributions of sound-meaning pairings in the lexicon, phonaesthemes are found to be pervasive in human languages. They have been documented in such diverse languages as English (Wallis 1699, Firth 1930, Marchand 1959, Bolinger 1949, 1980), Indonesian (McCune 1983) and other Austronesian languages (Blust 1988), Japanese (Hamano 1998), Ojibwa (Rhodes 1981), and Swedish (Abelin 1999). While it is difficult to establish reliable estimates of how many of the world's languages contain phonaesthemes, or how many words in a given language's lexicon have phonaesthemes in them, every systematic study of a particular language has produced results suggesting that that language has phonaesthemes. In general, phonaesthemes seem to appear in content words over function words, and in more specific (or subordinate level) rather than more general (or basic level) words.

In addition to distributional evidence, phonaesthemes can be detected through their role in language change, in particular by the part they play in the generation of neologisms (Joseph 1987). Phonaesthemes have been implicated in both the production and perception of neologisms on the basis of experimental (Abelin 1999) and comparative studies (Blust 1988, 2003). The recurrent finding of note in this latter area is that in a given set of related languages, phonaesthemes that appear in some languages will also appear in other languages but in words that are not cognates. For example, the Austronesian phonaestheme η occurs in initial position in words that have meanings related to 'nose' or 'mouth' (Blust 2003). In particular, it appears in words meaning 'chew, masticate, ruminate', such as Amis *ɲafɲaf*, Toba Batak *ɲaltok*, and Trukese *ɲú*, among others, but none of these words are cognates.

* Thanks to Nancy Chang, Jerome Feldman, Larry Hyman, Sharon Inkelas, George Lakoff, and Madelaine Plauché for helpful comments on drafts of related work, and to Amy Schafer and Robert Blust, in addition to two anonymous referees, for very useful suggestions on this paper.

While statistically significant representation in the lexicon and participation in the construction of neologisms are important convergent indications that phonaesthemes are a viable part of language, these two types of evidence leave room for different interpretations of what exactly the status is of phonaesthemes. The extent to which phonaesthemes play a role in the synchronic mental organization of language remains an open question.

Indeed, their cognitive status remains controversial in part because they do not fit well into linguistic theories that view compositionality as a defining characteristic of morphological complexity, such as so-called ITEM-AND-ARRANGEMENT models (Hockett 1954). On these compositional types of account, phonaesthemes belong outside the scope of regular morphological analysis. By contrast, noncompositional theories of morphology predict that recurrent sound-meaning pairings in a language, whether compositional or not, will rise to the status of organizing structures in the lexicon, on the basis of their frequency. Examples of such noncompositional models are the DYNAMIC MODEL (Bybee 1985), the USAGE-BASED MODEL (Langacker 1991), SEAMLESS MORPHOLOGY (Starosta 2003), and most connectionist morphological models (e.g. Plaut & Gonnerman 2000).

This article investigates the psychological reality of phonaesthemes, thereby providing an indication of the appropriate place for phonaesthemes in theories of morphology. Presenting language users with a sequential pair of words sharing a phonaestheme produces a processing advantage for the second word (a facilitatory priming effect). This phonaesthemic priming effect differs significantly from form priming, shown by words sharing only a phonological onset, and from semantic priming, which occurs between semantically related words. Rather, phonaesthemic priming very closely mirrors the priming effects that have been reported for morphologically related words. Moreover, phonaesthemic priming is not observed between simply any two words that by chance share both some phonological form and some meaning—it surfaces only when the form-meaning pairing is well attested in the lexicon. I conclude from these results that while compositionality may play a role in morphological organization, statistical prevalence in the lexicon of form-meaning pairings is also a sufficient criterion for those pairings to display morpheme-like behavior in language-processing tasks. As such, these results also lend support to the prediction made by noncompositional morphological models that the frequency, and not just the compositionality, of recurring form-meaning pairings is crucial to their mental representation.

2. PHONAESTHEMES AND COMPOSITIONALITY.

2.1. COMPOSITIONALITY IN MORPHOLOGICAL THEORIES. Human language is frequently characterized as a finite system with infinite capacity (e.g. Chomsky 1995). On most views, the key property of linguistic systems that gives rise to this infinite capacity is the possibility of combining linguistic units with others in novel ways. Compositional morphological theories, such as item-and-arrangement models, view this property, PRODUCTIVITY, to be predicated upon the hierarchical nature of certain linguistic structures. Complex linguistic units such as complex words are believed to be composed of smaller units, that is, they are COMPOSITIONAL. Most linguistic theories view compositionality, along with the resulting productivity, as a defining feature of human language (Janssen 1997, Partee 1995). Syntactic, semantic, phonological, and morphological compositionality have been particularly central in shaping mainstream linguistics of the second half of the twentieth century (Halle & Marantz 1993, Janssen 1997). The centrality of this capacity in theories of morphology is entirely unsurprising, since a large part of what

language users know about the words of their language is indeed how to combine morphemes to construct novel words.

Many morphological theories view complex words as compiled in production or decomposed in perception. The processes can be seen as the simple concatenation of morphological units by item-and-arrangement models (e.g. Lieber 1980, Di Sciullo & Williams 1987, and Halle & Marantz 1993).¹ What these models share is the view that the morphological capacity is something akin to a metaphorical assembly line, which either compiles a complex word from its morphological parts in production, or takes apart a complex word to analyze the pieces during perception. Moreover, for many theories of the word-internal relation between sound and meaning, the ability to combine units in novel ways to produce larger ones even defines the domain of study. On one view, for example, 'processes of word-formation that are no longer productive . . . are of little or no interest to morphological theory' (Baayen & Lieber 1991:802).

By itself, though, compositionality makes up only part of what language users know about how the words of their language pair form and meaning. They also have unconscious knowledge of groups of words that share meaning and are phonologically or prosodically similar (Serenio 1994, Kelly 1992). They know subtle correlations between a person's gender and the phonology of their name (Cassidy et al. 1999) and between the number of syllables in a word and the complexity of the object it describes (Kelly et al. 1990). They also know when the pieces of a word bear meaning, whether or not those parts can be productively combined with others (Bybee & Scheibman 1999).

As the studies cited above indicate, and as shown below, compositionality, as it is usually understood, does not suffice to capture what language users know about word-internal form-meaning relations. The next section describes the difficulties that compositional models face when confronted with phonaesthemes.

2.2. THE TROUBLE WITH PHONAESTHEMES. Phonaesthemes hold a prominent place among morphological phenomena that are troublesome to compositional accounts of morphology. Among these difficult cases are CRANBERRY MORPHS (Bloomfield 1933), like *boysen-*, which cannot stand alone and are permitted to attach to only one stem.² These morphs raise problems: what is the constituent structure of the words in which they occur, and what contribution (if any) do the cranberry morphs make to the meaning of the words as a whole? Phonaesthemes are similar in that, like cranberry morphs and bound morphs such as *-mit* in *transmit* and *submit*, they cannot stand on their own. And like other bound morphs but unlike cranberry morphs, phonaesthemes by definition appear in numerous words. Uniquely, though, phonaesthemes occur in words in which the remainder is often less morpheme-like than the phonaesthemes are themselves. Removing the phonaestheme *gl-* from words like *glint* and *glitter* leaves material that is not morphemic in nature or is at best a cranberry morph itself (though see the contrasting account in Rhodes & Lawler 1981). The combination of these last two properties of phonaesthemes—their frequency despite the fact that they combine with material of a questionable morphological nature—is what makes them particularly irksome.

¹ Word-and-paradigm and item-and-process models (Hockett 1954) do not make strong use of compositionality, though they do depend on productivity.

² *Cran* might be considered to no longer be a cranberry morph, as it has come in recent years to be used more extensively in compounds like *cran-apple*. In such compounds, however, *cran-* means 'cranberry' and thus may be different from *cran-* in *cranberry*.

Phonaesthemes are defined as form-meaning pairings that crucially are better attested in the lexicon of a language than would be predicted, all other things being equal. For an example of this type of hyper-attestation, consider the distribution of four phonaesthemes in the Brown Corpus (Francis & Kucera 1982). A full 39% of the word types and 60% of word tokens starting with *gl-* have definitions that relate to 'light' or 'vision' (Table 1) in an online version of *Webster's 7th collegiate dictionary*.³ Similarly, 28% of word types and 19% of word tokens with a *sn-* onset have meanings related to 'nose' or 'mouth'. Complete lists of each set of word types can be found in Appendix A.

	<i>gl-</i>		<i>sn-</i>		<i>sm-</i>		<i>fl-</i>	
	TYPES	TOKENS	TYPES	TOKENS	TYPES	TOKENS	TYPES	TOKENS
'light/ vision'	38.7% (48/124)	59.8% (363/607)	1.1% (1/8)	0.3% (1/31)	0% (0/7)	0% (0/1042)	10.9% (28/256)	10.1% (146/1441)
'nose/ mouth'	4% (5/12)	1.4% (9/607)	28.4% (25/88)	19.0% (59/311)	25.3% (19/75)	27.3% (284/1042)	5.6% (15/256)	2.7% (39/1441)

TABLE 1. Phonaesthemes *gl-*, *sn-*, *sm-*, and *fl-* in the Brown Corpus.

The overwhelming statistical pairings of forms like *gl-* and *sn-* with their associated meanings run contrary to what one would predict if the only correlations between the sounds of words and their meanings derive from shared component subparts, that is, constituent morphemes. On the basis of the assumption that a simple word's phonological form is entirely arbitrary, given its semantics, there should be the same portion of *gl-* words that have meanings related to 'light' or 'vision' as there are *sn-* words that share these meanings. And yet, as Table 1 shows, this is clearly not the case. While more than one third of *gl-* word types relate to 'light' or 'vision', only one percent of *sn-* words do. The reverse is true for *sn-* words and the meanings 'nose' and 'mouth'—while nearly one third of *sn-* word types relate to 'nose' or 'mouth', only four percent of *gl-* word types share this semantics. One can also compare the proportion of *gl-* or *sn-* words that bear these particular meanings with the distribution of *fl-* initial words in Table 1. While some *fl-* initial words have 'light' or 'vision' meanings, like *flicker* and *fluorescent*, these words are less numerous than are *gl-* words with similar meanings.

Since morphological models are attempts to understand distributions of lexical and sublexical sound-meaning pairings in a language, phonaesthemes appear at first blush to be perfect fodder for them. But phonaesthemes are problematic for compositional morphologies because words that contain these phonaesthemes also contain a complement that is usually not itself a meaningful unit. Removing the onset phonaesthemes from *glint* and *snarl*, for example, yields *-int* and *-arl*. It would be problematic both semantically and structurally to say that *-int* and *-arl* were units that contributed to the words as a whole. Neither do *-int* and *-arl* have meanings on their own, and we cannot mix and match the phonaesthemes and their complements: *glarl* and *snint* are mostly nonsensical (though as potential neologisms we can assign possible meanings to them, as discussed in §2.3 below).

It should be noted that although all the examples seen thus far are drawn from onsets, other parts of syllables, like rimes in words like English *smash*, *bash*, *crash*, and *mash* (Rhodes & Lawler 1981, Rhodes 1994) and even entire syllables in languages like

³ The distinction between word types, the list of word entries that appear in the corpus, and word tokens, the number of actual instances of each of those words, is a common and necessary one in studies of corpus frequency.

Indonesian (Blust 1988), can also be appropriately classified as phonaesthemes. Thus in some cases, a word may actually appear to be constructed compositionally entirely from phonaesthetic material. For example, *sneer* might be seen as composed from the onset phonaestheme *sn-* plus a rime phonaestheme *-eer*, also found in *leer* and *jeer*, with a meaning pertaining to 'expression of contempt'. Nevertheless, the overwhelming majority of words containing phonaesthemes are not compositional, as a perusal of the examples in Appendix A demonstrates. While phonaesthemes display some potential compositionality, this is not their usual behavior.

It is the tension between compositional models' apparent need to account for phonaesthemes and their inability to do so that drives us to look further into the details of the psychological status of phonaesthemes.

2.3. EXPERIMENTAL EVIDENCE ON PHONAESTHEMES. Phonaesthemes, as noted above, are pervasive in human languages. Traditionally, they have been documented primarily on the basis of distributional data. While there is a long tradition of psycholinguistic experiments focusing on how language users interact with novel words containing particular purported sound-symbolic elements such as /i/ and /a/ (see the summary in Jakobson & Waugh 1979), only recently has the psychological reality of large numbers of phonaesthemes been placed under systematic experimental scrutiny. As the reader will see, the psychological status of phonaesthemes has been difficult to pin down for methodological reasons.

The standard linguistic test for evaluating the psychological reality of a proposed linguistic unit is based on compositionality and productivity. In her seminal study of regular English plural marking, Berko (1958) argued for the psychological reality of morphological knowledge on the basis of subjects' generalizations of an existing pattern in the language to new forms, such as *wug*. When subjects accurately pluralized the novel form *wug* [wʌg] as *wugs* [wʌgz], this was taken as evidence that they had internalized a rule for forming plural nouns. Such neologism tasks (now known as WUG-TESTS) are pervasive in psychological studies of morphological knowledge.

But a methodological predisposition towards compositionality and productivity leads to the following quandary. It is quite difficult to assess the psychological reality of phonaesthemes, which are generally not compositional or very productive, if the only metric we have is crucially based on productivity and compositionality themselves. A partial solution to this dilemma has arisen from the observation that phonaesthemes, while noncompositional, are in fact partially productive.

In the past few years, three theses (Hutchins 1998 and Magnus 2000 on English, and Abelin 1999 on Swedish) have addressed the psychological reality of phonaesthemes from essentially the same perspective: that of neologisms. Magnus's work is representative of all three, so her thesis can serve as a model to demonstrate their methodologies. To test the psychological status of phonaesthemes Magnus uses two types of experiments based upon subjects' recognition and production of neologisms. The first method tests for knowledge of phonaesthemes by providing subjects with definitions for nonexistent words and asking them to invent new words for those definitions. In Magnus's study, subjects tended to invent novel words that made use of phonaesthemes of their language. Given the definition 'to scrape the black stuff off overdone toast', 27% of subjects invented a word that started with *sk-*. The second type of experiment tested the use of phonaesthemes in the perception of novel words. Subjects were presented a nonword and were asked to provide a definition for it. Again, they responded as if they were using phonaesthemes. For example, *glon* (bearing a *gl-*onset) evoked definitions relating to 'light' in 25% of the cases.

These studies each provide independent evidence for the role of phonaesthemes in the processing of neologisms. Magnus (2000) interprets her results as evidence that language users represent individual phones as bearing meaning (note that this is similar to Joseph's (1994) conclusion for the Greek phoneme /ts/, made on the basis of distributional evidence). Abelin (1999) adds to this the insight that the fewer words there are that share the phonology of a phonaestheme (excluding the phonaesthetic words themselves), and the more common that phonaestheme is among the words sharing its phonology, the more likely subjects are to treat neologisms along the lines of that phonaestheme. Hutchins's (1998) results confirm the role of phonaesthemes, even crosslinguistic ones, in the processing of neologisms. Interestingly, these findings are qualitatively similar to results of neologism experiments on purely phonological generalizations, such as the obligatory contour principle in Arabic (Frisch & Zawaydeh 2001) and labial attraction in Turkish (Zimmer 1969)—frequent patterns, whether phonological or phonological-semantic, are used in neologism interpretation and production.

Probably the most famous use of English language literary neologisms are in 'The Jabberwocky', a poem that confounds young Alice in Lewis Carroll's (1897) *Through the looking glass*. The first stanza of this poem appears in 2.

- (2) 'Twas brillig, and the slithy toves
 Did gyre and gimble in the wabe:
 All mimsy were the borogoves,
 And the mome raths outgrabe.

Luckily for Alice, Humpty Dumpty is gracious enough to explain the poem: 'Well, "slithy" means "lithe and slimy"' (Carroll 1897) and so on. This literary example is particularly illustrative of the problem with neologism studies.

Although they provide valuable insight into how language users interact with new words, neologism studies do not by themselves constitute conclusive evidence for the psychological status of phonaesthemes. Even in light of the neologism results from the three phonaestheme works cited above, one could still hold the position that phonaesthemes are only static, distributional facts about the lexicon, which speakers of a language can access consciously. This is problematic since essentially all normal morphological processing happens unconsciously. We know that language users are able to access all sorts of facts about their language upon reflection. People can come up with a word of their language that is spelled with all five vowel letters and 'y' in order, or a word that has three sets of double letters in a row.⁴ These abilities by themselves, though, do not lead to the conclusion that orthographic order of vowel letters in a word is a fundamental principle of implicit cognitive organization. For the same reason, subjects' ability to consciously access distributions of sound-meaning pairings in their language does not imply that those pairings are meaningful for the subjects' linguistic system.

In fact, different mechanisms could be evoked to account for the cognitive processing that the neologism studies described above demonstrate.

- a. Subjects could be taking a sample of the lexicon, and observing generalizations over a set of words that share structure with the neologisms.
- b. They might be randomly selecting nearest neighbors.

⁴ *Facetiously* and *bookkeeper*, respectively, although Brian Joseph (p.c.) notes that if one is willing to adopt a flexible definition of *word*, then a word with five double sets of letters in sequence is *boobbookkeeper* (a bookkeeper who is a boob).

- c. They could be looking for cues in the stimulus itself; notice that the word ‘scrape’ appears in the example for *sk-* provided above, and the word ‘slimy’ appears in Humpty Dumpty’s explanation for ‘slithy’.
- d. Finally, subjects just might be unconsciously acting in accordance with the structural configuration of their language system, which is structured in part on the basis of phonaesthemes.

Unfortunately, we cannot distinguish among these solutions on the basis of neologism tasks like the ones described above.

What is needed to resolve this issue, then, is a way to tap into unconscious language processing of morphemes. Applying such a measure to phonaesthemes, one can evaluate the extent to which their behavior in language processing is like or unlike concatenative, productive morphological units. To reiterate, the stakes are high. If phonaesthemes have a demonstrable psychological reality, then the notion of morphemes as concatenative, productive units is insufficient for defining the human capacity to extract and use generalizations about the internal structure of words.

3. METHOD. From the perspective of their distribution, phonaesthemes are statistically well-represented partial form-meaning pairings. In order to test whether phonaesthemes also have a psychological reality—whether they are internalized and used by speaker-hearers—we have to answer two related questions. First, does the presence of a phonaestheme in a word affect the processing of that word? If phonaesthemes play a role in language processing, then words in which they occur should be processed differently from words in which they do not. Second, if phonaesthemes do in fact affect lexical processing, is this effect significantly different for phonaesthemes than for subparts of words that correlate some form and some meaning, but not in a statistically significant way in the lexicon? In other words, is there a role for the frequency of a form-meaning pairing in determining processing effects?

The morphological priming methodology (first used by Kempley and Morton (1982), summarized by Drews (1996)) provides an excellent starting point for developing a method for addressing such empirical questions. Morphological priming is the facilitation (speeding up) or inhibition (slowing down) of mental access to a TARGET word on the basis of some other PRIME word, which has been presented previously. A large number of morphological priming studies have found priming patterns that are unique to morphologically related prime and target words, and are not shared by morphologically unrelated words (see the summary in Feldman 1995). This difference between morphological priming effects on the one hand and known phonological and semantic priming effects on the other has been used to argue that morphemes have a psychological reality. In the same way, we can hypothesize that if phonaesthemes have some cognitive status, then they should similarly display priming effects that are different from those exhibited by words that do not share phonaesthemes. The study described below aimed to test this possibility.

In this study, twenty native speakers of English, aged 18 to 36, were first presented with a prime stimulus, which was an orthographic representation of a word, on the screen of a digital computer. Following Feldman and Soltano’s setup (1999), it appeared slightly above the center of the screen for 150ms, just long enough to be barely perceived by the subject. Three hundred milliseconds later, a second stimulus, the target stimulus, also a typewritten word, was presented in the center of the screen for 1000ms or until the subject responded to it. This inter-stimulus latency of 300ms was chosen because that delay has been reported to most effectively separate morphological from nonmor-

phonological priming effects (Feldman & Soltano 1999). At this inter-stimulus interval, there is the least phonological or semantic priming relative to morphological priming. Subjects performed a lexical decision task; they were asked to decide as quickly as possible whether the target was a word of English or not. They were to indicate their decision by pressing one of two keys on the computer keyboard—z for nonwords and m for words.

The critical stimuli were fifty pairs of words, falling into five categories (Table 2). In the first condition, the prime and target shared some phonological feature, always a complex onset, and some semantic feature, such as ‘nose’ or ‘light’. Additionally, a significant proportion of the words in English sharing that onset also shared the semantics, giving pairs like *glitter* and *glow*. In the second condition, the prime and target shared only a phonological feature, always an onset, yielding stimuli such as *druid* and *drip*. Primes and targets in the third condition shared only some semantics, like *cord* and *rope*. In the fourth condition, the pseudo-phonaestheme case, stimuli shared phonology and semantics but were part of a very small number of words occurring in the Brown Corpus (usually only two) that did so. An example of such a pair is the nouns *crony* and *crook*. Finally, for comparison, stimuli in the fifth, baseline condition did not share any segments in their onset, or a meaning, giving pairs like *frill* and *barn*. The full list of stimuli presented can be seen in Appendix B.

Since the subjects were performing a lexical decision task, choosing whether or not a string of characters was a word of English, an equal number of nonwords was required as well. An additional fifty nonce words were thus created by manipulating one or two characters in each of the existing target words (also seen in Appendix B). For example, the real target word *glow* became *glone* and *crook* became *croof*. Crucially, all of these nonwords shared their onset (or in the case of the one vowel-initial word, the initial vowel) with the real target word from which they were manufactured. The nonce target words were paired with the prime words that the words that were manipulated to form them were paired with. This ensured that subjects could not simply rely on a shared onset between prime and target, or indeed a given onset on a target word (such as *sn-* or *gl-*) as a cue to whether the target was a word. Each subject saw half of the prime words in each condition followed by a real-word target and the other half followed by a manipulated nonword target.

CONDITION	CHARACTERISTICS	EXAMPLES
PHONAESTHEME	Prime and target share a semantic feature and phonological onset; are a statistically significant subclass of the lexicon	<i>glitter:glow</i>
FORM	Prime and target share an onset	<i>druid:drip</i>
MEANING	Prime and target share a semantic feature	<i>cord:rope</i>
PSEUDO-PHONAESTHEME	Prime and target share a semantic feature and an onset; are not a statistically significant subclass of the lexicon	<i>crony:crook</i>
BASELINE	Prime and target unrelated in form and meaning	<i>frill:barn</i>

TABLE 2. The five test conditions.

Translated into specific hypotheses about the results of this study, the questions at the beginning of this section appear as follows. First, if the presence of phonaesthemes affects processing, then responses to condition 1 (the phonaestheme case) should be significantly different from those to condition 5 (the baseline, unrelated condition). To be certain that this effect does not derive from known priming effects resulting from phonological (Zwitserslood 1996) and semantic relatedness (Thompson-Schill et al. 1998) between prime and target, the responses to class 1 must also be distinct from those to classes 2 and 3. Second, if this priming effect is restricted to statistically well-

represented form-meaning pairs, then responses to condition 1 (phonaesthemes) should be significantly different from those to condition 4 (pseudo-phonaesthemes).

Before assessing the results of the experiment, I must briefly discuss some controls applied to it. The morphological priming methodology assumes that it takes subjects the same length of time to make lexical decisions for the targets in each of the different conditions. If a set of target words in one condition requires a longer lexical decision than those in another condition, then any differences in response time across conditions could be a result of this difference, rather than different priming effects. In fact, in most experiments using this methodology, the same targets are used across conditions, with only the primes varying. The current experimental design requires different target words in the different conditions, as there are very few words (if any) that simultaneously instantiate phonaesthemes and pseudo-phonaesthemes. Two important factors that influence the rate of decision to a given target word—length and frequency—must therefore be controlled for.

The target words in all conditions were matched as closely as possible for average token frequency and length, since these factors have demonstrable effects on response time and morphological priming (Meunier & Segui 1999). As seen in Table 3, it was possible to find stimuli whose lengths were, on average, very similar—no conditions differ by as much as a phone or letter. Frequency is much more difficult to assess, since the type of corpus one selects for frequency measures radically influences the rates of occurrence of the words contained in it. Table 3 shows two standard written frequency measures, Thorndike-Lorge and Kucera-Francis, both taken from the MRC Psycholinguistic database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). While there is good consistency across most of the conditions for each of these measures, it is clear that according to Thorndike-Lorge, target words in the baseline and pseudo-phonaestheme conditions are the least frequent, while the Kucera-Francis numbers place the pseudo-phonaestheme words as least frequent.

	Thorndike- Lorge	Kucera- Francis	LETTERS	PHONES	SYLLABLES
PHONAESTHEME	181.4	15.75	5.2	4	1
FORM	136.56	12.7	5	4.5	1.1
MEANING	160.5	13.1	4.7	4.1	1.4
PSEUDO-PHONAESTHEME	111.5	5.5	5.1	4.5	1.1
BASELINE	96.3	12.22	4.8	3.8	1.2

TABLE 3. For each condition, mean frequency (written frequency in Thorndike-Lorge and Kucera-Francis), and length (in letters, phones, and syllables).

These frequency differences should be a matter of concern, since our main interest is in determining whether different types of relations between a prime and target yield different degrees of priming. It might take longer to make a lexical decision about less frequent target words, such as those in the pseudo-phonaestheme condition, regardless of what precedes them, in which case it would be impossible to determine whether any differences among the conditions resulted from effects of the prime stimuli or simply from different default rates of response to the target words.

Fortunately, there exist measures of lexical decision times for words in isolation, which help us determine how serious the frequency differences described above are. Table 4 shows the mean response time in a lexical decision task to the target words in each of the five conditions, according to the English Lexicon Project (Balota et al. 2002). As we can see, the phonaestheme, form, and meaning conditions are quite similar, while target words in the pseudo-phonaestheme condition are responded to

more quickly and those in the baseline condition more slowly. These differences are taken into consideration in analyzing the results from this experiment in §4.

	MEAN RT IN ISOLATION
PHONAESTHEME	645
MEANING	631
FORM	655
PSEUDO-PHONAESTHEME	610
BASELINE	681

TABLE 4. For each condition, mean reaction time to the stimuli in isolation (according to the English Lexicon Project).

4. RESULTS. The average reaction times per condition fell out as shown in the first row of Table 5. By far, the fastest average reaction time was to the phonaestheme condition, which was 59ms faster than the baseline (unrelated) condition. That is to say, subjects responded to a target stimulus 59ms faster when the prime and target shared a phonaestheme, relative to the condition where prime and target were semantically and phonologically unrelated. By comparison, prime and target pairs that shared meaning but no form were identified only 23ms faster than the baseline condition. By contrast, words sharing an onset with their prime, those in the form condition, took slightly longer to decide on than did those in the baseline condition. In other words, phonaesthemically related words led to much faster recognition of the target than did semantically or phonologically related words.

	PHONAESTHEME	FORM	MEANING	PSEUDO-PHONAESTHEME	BASELINE
Average RT under priming (ms)	606.7	668.2	642.7	658.7	665.3
Average RT in isolation (ms)	645	655	631	610	681
Difference (ms)	-38.3	13.2	11.7	48.7	-15.7

TABLE 5. Average target reaction times by condition, mean reaction time to the stimuli in isolation (according to the English Lexicon project), and the difference.

Moreover, shared form and meaning together were not sufficient to explain the phonaesthemic priming. Recall that the phonaestheme condition elicited reaction times 59ms faster than the baseline. By comparison, the pseudo-phonaestheme condition, containing words that shared form and meaning but were among very few words in the lexicon sharing both, yielded reactions only 7ms faster than the baseline. This difference indicates that the extent of the form-meaning pairing's distribution in the language is crucial to its role in processing. Statistically prevalent form-meaning pairings yield faster responses than do similar words that are statistical loners. The significance of the effect of condition on reaction time was confirmed by a one-way ANOVA subject analysis ($F(4,95) = 3.76, p < 0.01$). The differences between the phonaestheme condition and each other condition were also evaluated using Fisher's PLSD and Sheffe's post hoc tests (statistical tests that evaluate the significance of the pairwise differences among levels of a conditioning factor), which both confirmed that the phonaestheme condition was significantly different ($p < 0.05$) from the four other conditions.

These results very clearly indicate that targets are responded to much more quickly when they share a phonaestheme with their prime than when they share form, meaning, both, or nothing with it. But as mentioned in the previous section, there remains the possible confound of how long it takes to make a lexical decision on the targets in the different conditions in isolation. If the reaction time differences we observed in the

first row of Table 5 are due solely to different decision rates for the targets in the different conditions, then there should be a strong correlation between the default response time and the primed response time, which might explain away the differences among the conditions. As might be hypothesized from a scan of the data in Table 5, there is no such statistically significant correlation ($r(5) = 0.2, p = 0.77$).

Giving these data a more detailed analysis, one might be interested in investigating whether differences in the reaction time between the phonaestheme condition and any others are in fact due to absolute differences in response time. One way to assess this is to compare the average difference between reaction time in isolation and in the priming experiment for each condition. If the phonaestheme condition is responded to faster under priming than in isolation, while the other conditions are not, then this is yet another indication that it is the phonaesthetic relation, rather than independent response time, that is responsible for the fast decisions in the phonaesthetic condition reported above. This is almost precisely what we find: as seen in the last row of Table 5, targets in the phonaestheme are responded to much faster (38.3ms faster) in the current priming experiment than in isolation, while the form, meaning, and pseudo-phonaestheme targets require slower lexical decision responses in the priming experiment than in isolation. The baseline condition showed a slightly faster response in the priming experiment than in isolation. Since the actual conditions of the current experiment and those that yielded the mean isolated lexical decision times differ, their results cannot be quantitatively compared. Nevertheless, they support the conclusion that even when the response times to these target words in isolation are taken into account, the faster response times in the phonaestheme condition do not result from shared form or meaning, or both. The role of reaction time in isolation was tested quantitatively by re-running the ANOVA described above, this time including reaction time in isolation as a covariate, making the new analysis an ANCOVA. In this analysis, the effect of the relation between the two words on reaction time was still significant ($F(3,95) = 4.82, p < 0.01$).

One final confound should also be addressed—the possibility that the stimuli in the phonaestheme condition are more closely related in terms of meaning than those in the other conditions, and that this is the cause for faster response time to words in this condition. While it is difficult to establish reliable and useful semantic similarity ratings, it is clearly useful to try to apply some quantitative analysis to the problem. One useful semantic similarity metric is a similarity rating produced by LATENT SEMANTIC ANALYSIS (LSA; see Landauer et al. 1998). LSA is, among other things, a statistical method for extracting and representing the similarity between words or texts on the basis of the contexts in which they appear. Two words or texts are rated as more similar the more alike their distributions are. LSA has been shown to perform quite like humans in a range of behaviors, including synonym and multiple-choice tasks. Of relevance to the current discussion is the pairwise comparison function, which produces a similarity rating from -1 to 1 for any pair of texts. Identical texts have a rating of 1 , while completely dissimilar ones would have a rating of -1 . To give examples from the pairs used in the phonaestheme experiment, *glitter* and *glow* have an LSA value of 0.27 (somewhat similar), while *barn* and *frill* have an LSA of 0 (no particular similarity or dissimilarity).

The mean LSA scores for each condition were as seen in Table 6. While the LSA values for word pairs in the form and baseline conditions were close to zero, those of the pseudo-phonaestheme, phonaestheme, and meaning conditions were higher. The significance of differences among these conditions was evaluated using an ANOVA,

which took the semantic relatedness of each prime-target pair, measured as an LSA value, as the dependent variable, and the condition (phonaestheme, form, meaning, etc.) as the independent variable. There was a significant main effect of condition ($F(4,42) = 6.02, p < 0.001$). According to Fischer's PLSD (a liberal test) and Sheffe's (a conservative one) post hoc analyses, the phonaestheme condition was not significantly different from the pseudo-phonaestheme condition (LSD $p = 0.07$; Sheffe $p = 0.51$), or from the meaning condition (LSD $p = 0.55$; Sheffe $p = 0.98$) but was marginally significantly different from the form (LSD $p < 0.01$; Sheffe $p = 0.08$) and baseline conditions (LSD $p < 0.01$; Sheffe $p < 0.1$). In other words, the pairs of words in the phonaestheme condition were just as similar in terms of meaning as were the words in the meaning-only and pseudo-phonaestheme conditions, while word pairs in the form-only and baseline conditions were significantly different.

	Mean LSA rating for prime-target pairs
PHONAESTHEME	0.23
MEANING	0.26
FORM	0.06
PSEUDO-PHONAESTHEME	0.12
BASELINE	0.06

TABLE 6. For each condition, mean LSA semantic similarity rating for word pairs.

The current experiment has demonstrated phonaesthemically related facilitatory priming by primes that are both semantically and phonologically related to a target that cannot be accounted for simply on the basis of other sorts of priming. The meaning priming and form priming both yield quantifiably different priming effects from phonaesthemic priming. As a potential complication, it has been proposed in the morphological priming literature that since stimuli that share a morpheme share both form and meaning, the sum of the priming effects from each of these domains might give rise to the actual priming observed with morphemes (see Feldman & Soltano 1999 for discussion). Taken together, however, form and meaning priming are still distinct from phonaestheme priming no matter how they are measured. If the difference between the baseline reaction time and the reaction time in every other condition is taken as the priming effect for that condition, then the sum of the meaning and form priming effects (20ms) differs a great deal from the phonaesthemic priming effect (59ms). If, however, it is the difference between the reaction time to a condition in isolation and the time to respond in the priming context that is used, then the difference is even greater—the form and meaning priming effects would sum to a delay of -25 ms, while the phonaesthemic priming effect would be an increase in speed at 38ms. No other metrics aside from summation have been proposed in the literature to deal with combined phonological and semantic effects. Phonaesthemic priming is an entity unto itself, resulting from the statistical over-representation of a particular pairing between form and meaning in the lexicon.

These results disconfirm the possibility that language users access phonaesthemic knowledge only during tasks allowing reflection, such as the neologism tasks described in §2.3 above. Instead, phonaestheme effects emerge even while individuals are performing a task that is tightly constrained by time pressure, and is therefore processed unconsciously, like natural language is.

5. DISCUSSION. The experimental evidence described above argues for the psychological reality of phonaesthemic priming effects. These effects bear implications for

questions as fundamental as the definition of the morpheme and the role of frequency in language learning, representation, and change.

5.1. PHONAESTHEME AND MORPHEME PRIMING. Despite the fact that they are mostly noncompositional, phonaesthemes behave just like central morphological units in terms of two definitional criteria. First, unlike the most frequent and productive affixes, but like other semiproductive units like English *en-* (Baayen & Lieber 1991), phonaesthemes are partially productive. Evidence of this productivity comes from the use of phonaesthemes in interactions with neologisms, as shown in Hutchins 1998 and Magnus 2000. Second, as the reader will now see, phonaesthemes mirror the priming behavior of canonical morphemes.

In general, morphologically related forms demonstrate facilitatory priming that differs from both semantic priming and phonological priming in terms of its degree and time course. Consider the results of an experiment described by Feldman and Soltano (1999), which was nearly identical in format to the one described above. The key difference between the two experiments was that for Feldman and Soltano, the test stimuli were morphologically related instead of being phonaesthemically related. Unlike the phonaestheme stimuli, the morphologically related forms in Feldman and Soltano's experiment shared a productive suffix, like the regular past tense. Table 7 lays out the priming effects from Feldman and Soltano's experiment in comparison with those from the phonaestheme experiment (measured as the difference between the mean reaction time of the baseline condition and that of each of the other conditions).

	RELATION BETWEEN PRIME AND TARGET			
	MORPHOLOGICAL	PHONAESTHEMIC	FORMAL	SEMANTIC
FELDMAN & SOLTANO	49	–	–18	34
PRESENT STUDY	–	59	–3	23

TABLE 7. Morphological priming and phonaesthemic priming in milliseconds.

The measures of priming in these two experiments cannot be compared quantitatively because the test conditions differed slightly, but a valuable conclusion can be drawn from their juxtaposition. The phonaesthemic priming detected in the experiment described above is qualitatively similar to evidence for morphological priming, like the morphological priming reported by Feldman and Soltano. In both cases, at the same inter-stimulus interval, the morphologically or phonaesthemically related condition yielded significantly greater facilitatory priming than form-only priming, meaning-only priming, or the sum of form and meaning priming.

This similarity between phonaesthemic priming and morphological priming, along with the partial productivity of phonaesthemes, suggests that phonaesthemes have a status in the mental language processing system that is similar to that of canonical morphemes. Observations of the importance of phonaesthemes as organizing units has led some, such as Blust (1988), to suggest that phonaesthemes may be represented as units, below the level of the morpheme, but above the level of the phoneme. While such a solution is enticing since it could preserve the compositionality of morphemes, it nevertheless spawns as many important difficulties. For one, words containing phonaesthemes are not compositional, which leaves open the question of how the remainder of a morpheme is represented at the phonaesthemic level. Second, instantiating a phonaesthemic level provides no a priori explanation for why frequent sound-meaning pairings would give rise to morpheme-like and not phoneme-like priming effects. Third, and perhaps most important, such a solution does not provide any explanation for the

importance of frequency effects at the phonaesthetic level. And phonaesthemes, as seen below, are just the tip of the frequency iceberg.

5.2. OTHER STUDIES OF STATISTICAL SOUND-MEANING PAIRINGS. Virtually any form-meaning pairing with sufficient statistical reliability can play a demonstrable role in the mental processing of language. The range of statistical form-meaning pairings that have been shown to play a role in processing (of which only a selection are surveyed below) testifies to the industriousness of the human capacity to encode statistical form-meaning correlations within words. Research on a number of seemingly unrelated topics indicates that language users integrate and make use of statistical correlations between sound and meaning, even when these relations do not play a productive role in the linguistic system.

Language users internalize subtle correlations between a person's gender and the phonology of their name. Cassidy et al. 1999 documented these correlations, finding that while male names in English like *Richard* and *Arthur* tend to have trochaic (word-initial) stress, female names like *Irene* and *Michelle* tend to have iambic (word-final) stress. In their corpus of 490 names, iambic disyllabic words strongly tended to be women's names (94%), while trochaic words did not (54% were female). Additionally, male names like *Bob* and *Ted* tend to end in consonants (with 80% of consonant-final words in their corpus being men's names), while female names like *Sue* and *Mary* tend to end in vowels (with 72% of vowel-final words in their corpus being women's names.) The important point in this work is that these correlations are simply statistical tendencies, and in some cases nearly as many counterexamples can be found as positive examples. The question was whether language users were aware of these correlations between a name's phonological form and the gender of the person to which it belongs. The researchers presented adults and four-year-old children with words that bore statistically prevalent characteristics of female or male first names, asking them to perform neologism and sentence completion tasks. Cassidy and her colleagues found that subjects tended to process the names along the lines predicted by the phonological cues to gender. Consonant-final words, for example, were deemed to be names for males 66% of the time by four-year-olds and 70% of the time by adults, both of which were significantly higher than their judgments for vowel-final words. What is interesting about these results is that they obtained despite the fact that the phonological cues to name gender are neither categorical nor compositional.

In a similar study, Kelly and colleagues (1990) found that adults and children have internalized and make use of a correlation in English between the number of syllables in a word and the complexity of the object it describes. Geometric shapes are particularly exemplary of this phenomenon; consider the tendency expressed by the correlation between the geometric complexity and syllabicity of *line*, *square*, and *point* with those of *trapezoid*, *hexagon*, and *parabola*. When presented with novel shapes, subjects were more likely to pair complex shapes with polysyllabic words than they were simpler shapes.

These findings are analogous to those described for phonaesthemes above—they provide external support that language users internalize statistical form-meaning correlations, despite noncompositionality, and make use of them during language processing. Similar results have been shown for other pairings between form and other variables, like a particular irregular past tense schema in English shared by verbs like *strung*, *hung*, and *slunk*, which pair the meaning 'past tense' with a vowel and phonological features of the onset and coda (Bybee & Slobin 1982, Bybee & Moder 1983). Another

example is the subtle but significant asymmetries in the distribution of phonological features across grammatical categories. Sereno (1994) demonstrated that frequent English verbs more often have front vowels than back vowels, while frequent nouns have more back vowels than front vowels, and that language users do indeed use knowledge of these asymmetrical phonosyntactic generalizations during perception. When asked to determine as quickly as possible whether a word they were presented with was a noun or a verb, subjects responded significantly faster when the word was a noun and had a front vowel than when it was a noun with a back vowel, and vice versa for verbs. Another well-known correlation between grammatical category and phonology is the statistical prevalence for disyllabic nouns in English to be trochaic and for disyllabic verbs to be iambic (Sherman 1975, Hayes 1980). Kelly and Bock (1988) showed that when pronouncing disyllabic nonwords in sentences that framed them as either verbs or nouns, subjects were significantly more likely to give the words initial stress if they were functioning as nouns than when they were supposed to be verbs. Kelly (1988) showed that when presented with a novel disyllabic word, subjects strongly tended to classify and use it as a noun if it had initial stress.

We can gather from these studies that language users unconsciously pick up on a broad range of form-meaning pairings and use them during language production and perception, even when the pairings are noncompositional and noncategorical. How exactly they do this is another question, one that can be answered in part by turning to probabilistic theories of linguistic knowledge and use (see for example the chapters in Bod et al. 2003). More generally, theories that are capable of accounting for the knowledge and use of phonaesthemes share the property of being grounded in the actual use of language.

5.3. USAGE-BASED MODELS. The key property that phonaesthemes and the three other cases summarized in §5.2 share is that their only manifestation is as recurrent pairings of formal and semantic properties of words. For phonaesthemes, this recurrence is sufficient to allow them to play a role in unconscious language processing, as well as, possibly, in the invention and interpretation of neologisms. While compositional morphologies have very little to say about the mechanisms by which phonaesthemes might be learned and represented, usage-based views of morphology predict the occurrence of phonaesthemes as a natural product of generalization over the lexicon. Network models (e.g. Bybee 1985, Langacker 1991) and connectionist models (e.g. Plaut & Gonnerman 2000) both predict that statistical recurrences across words, like phonaesthemes, will automatically rise to the status of organizing structures in a language.

Acquiring phonaesthemes, like all the other form-meaning pairings described above, is accounted for in network models as the emergent product of implicit abstraction over specific lexical representations. In Bybee's (1985) network morphology, for example, phonological and semantic similarities across words give rise to connections between them. This might be represented as in Figure 1, where we can see that connections between words (depicted as solid lines) emerge where form elements or meaning elements are shared across words. Where form and meaning are shared across a set of words, a morphological relation emerges; that is, a generalization appears where previously there were only memorized word forms. This emergent generalization is usually referred to as a schema (Bybee & Slobin 1982). Morphological schemata, as simple generalizations over individual tokens, are based not on compositionality or productivity, but rather on the simple recurrence in the language of a form-meaning association. Langacker's model (1991) works similarly, generalizing over individual examples such

that more abstract schemas emerge. Janda (1983) proposed a similar model, in which morphemes may overlap such that both *gl* in *glisten* and the entire word itself simultaneously instantiate different morphemes. In all three of these models, the only difference between compositional and noncompositional morphological units is whether the remainder of the words in which they occur also instantiates a more general schema.

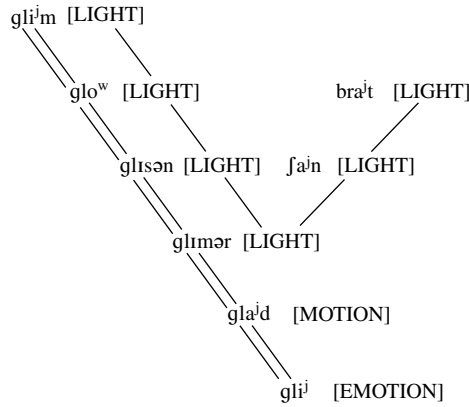


FIGURE 1. Piece of a network model of generalization over phonaesthemes.

Distributed connectionist models of morphology share the premise that morphological relations are emergent from semantic and phonological generalizations over stored lexical forms. They differ from network models like Bybee’s, however, in two main ways. First, they are quantitative models, a fact that facilitates their evaluation when they are implemented in a computational environment. Second, rather than a unique representation for each word, they often posit a network of phonological and semantic units, where a pattern of activation over those units represents a word. In the representative miniature connectionist model in Figure 2, a tendency for form units *g* and *l* to cooccur with the meaning unit ‘light’ is quantitatively encoded in the strength of connections among the three units. Activation of *g* and *l* increases the likelihood that ‘light’ will also become active, due to emergent strengthening of connections between these nodes and the network’s internal representation in its hidden layer.

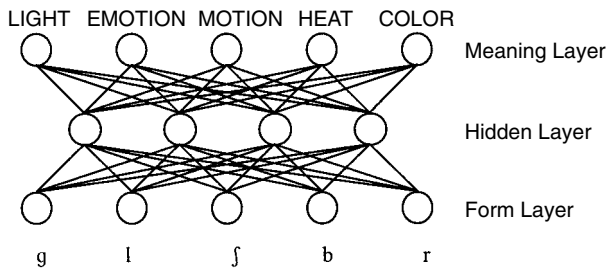


FIGURE 2. Miniature connectionist model for learning and representing phonaesthemes.

In all usage-based models, connectionist or otherwise, the strength of an emergent schema, which allows it to play a role in unconscious and conscious language tasks, depends at least on its frequency of occurrence in the words of a language and the frequency of those words that exemplify it. Schema strength does not depend directly

on the compositionality of the form-meaning pairing. Thus, usage-based morphologies provide a natural explanation for the behavior of phonaesthemes we have seen above. In a sense, usage-based models advocate a return to the classical view that a morpheme is at its core a recurrent sound-meaning correlation across the lexicon, as contrasted with item-and-arrangement models, which view morphemes as only those such correlated patterns that also participate in compositional arrangements. Where usage-based models move beyond this simple definition of the morpheme is in their emphasis on the central role of type and token frequency in the development of such morphological representations.

It is important to note that a certain class of morphological theories is constructed from two independent morphologies—an item-and-arrangement model and a usage-based model. These DUAL ROUTE models (introduced in Pinker & Prince 1988) recognize that not all morphology is compositional and view the morphological capacity as composed of a regular, compositional component and an irregular, lexical component. Phonaesthemes would in such a model be placed in this second group. But separating morphological knowledge into two separate modules raises a host of problems, particularly when phonaesthemes are on the table. In the discussion above, the only demonstrated difference between phonaesthemes and other morphological units is their compositionality (but see Rhodes & Lawler 1981). They can all be more or less productive and can be used to invent and interpret neologisms, and they demonstrate priming effects that are distinct from phonological and semantic priming effects. Thus, a dual route model would need to explain why all these properties are shared by compositional and noncompositional morphological units, which are processed in two separate routes. This is a less parsimonious solution than those, like usage-based models, that use a single mechanism for both types of morphological unit.

5.4. FUTURE DIRECTIONS. The present work raises the question of the precise relationship between morphological and phonaesthetic priming. Only by combining these two types of test stimuli in a single experiment can we answer this question. Also of future interest is the specific role that statistical frequency plays in phonaesthetic priming. Meunier and Segui (1999) have shown that frequency plays a role in morphological priming, but we do not yet know what the exact character of the effect of frequency in phonaesthetic priming is. Is being a phonaestheme an all-or-none enterprise, or does frequency correlate gradedly with degree of phonaesthetic priming? The present study did not include frequency as a graded parameter but rather as a discrete one; statistically significant phonaesthemes were assigned to the same condition, regardless of how significant they were. Another factor that might play a role in phonaesthetic priming is the degree of semantic overlap between prime and target. This factor has been shown to play a role in the degree of morphological priming between stimuli and might also do so for phonaesthemes (Laura Gonnerman, p.c.).

In addressing the psychological reality of phonaesthemes, I have addressed only the degree to which phonaesthetic knowledge has been internalized by language users. Other questions about phonaesthemes deserve attention, though. The study described above assumes that the potential universality of or motivation for phonaesthemes is orthogonal to their representation. In contrast with this approach, phonaesthemes have predominantly been investigated not in terms of their representation, but in terms of their motivation—that is, with respect to their sound-symbolic properties. Various studies have attempted to document crosslinguistic phonaesthemes for the purpose of demonstrating that there is some correlation between the sound structure of words that have

phonaesthemes and the meanings they bear. Famous cases include the often reported use of high vowels to express small size and low vowels to indicate largeness, as in English *teeny* and *tiny* versus *humungous* and *large* (e.g. Ohala 1984). The results of these studies are frequently interpreted as disproving the arbitrariness of the sign (e.g. Bolinger 1949). Both the results and this conclusion, though, are hotly contested. While motivation usually describes ultimate historical or acquisitional causes for the phonaestheme's existence, there may in fact also be some relationship between the semantic motivation for a phonaestheme and its distribution. More semantically unified or more semantically 'basic' meanings, however basicness is measured, might be more prevalent. One way to test these hypotheses would be to repeat the experiment presented above with phonaestheme stimuli that varied in terms of their proposed universality.

And finally, if statistical prevalence of a sound-meaning pairing in a language results in a demonstrable representation of that pairing in the cognitive system, which itself may give rise to the ability of individuals to invent new words based on that pairing and to interpret others' neologisms that make use of that pairing, then we would expect the statistical distribution of a phonaestheme to affect its extension in a language over time. Phonaesthemes that are well represented should yield new forms over time for which there is no cognate in closely related languages. One such case, discussed by Hock and Joseph (1996:293), concerns the rime of words like *drag*, *flag*, *lag*, and *sag*, which all share the meaning of 'slow, tiring, tedious motion'. According to Hock and Joseph, this cluster gained a new member in the sixteenth century when English *sacke* became *sag* through an irregular, analogical change, which they name PHONESTHEMATIC ATTRACTION. Elsewhere, Blust (1988) reports on the historical developments of Austronesian roots, which are like English phonaesthemes except that they occur word finally and can encompass as much as an entire syllable. Blust has found numerous cases of cognateless forms in various Austronesian languages which, it seems, are best explained as cases of the richer roots getting richer, or as Joseph (1983:ch. 7) and Blust (2003) call it, a 'snowballing effect'. In the future, I hope to investigate how much of a role the precise distributional statistics of phonaesthemes play in influencing their extension over time.

6. CONCLUSION. Like other noncategorical pairings between phonology and semantics, phonaesthemes have a significant psychological status. Specifically, when a form-meaning pairing recurs sufficiently often, it comes to take on priming behavior that cannot be explained as the result of form or meaning priming, alone or in combination. Item-and-arrangement and dual route models of morphology are unable to account for the similarity between phonaesthetic priming and morphological priming, and the experimental results have been interpreted as supporting usage-based models. Phonaesthemes are a testament to the diligence of the human ability to encode and use subtle statistical associations in the linguistic environment.

APPENDIX A

These words were among those extracted from the Brown Corpus. They were selected as containing the phonaesthetic meanings if one of their senses in Webster's 7th dictionary referred to some aspect of 'light' or 'vision' for the case of *gl-* and 'nose' or 'mouth' for *sn-*.

gl-

glance
glanced
glances
glancing
glare

glimmering
glimpse
glimpsed
glimpses
glint

sn-

snack
snacks
snarled
snarling
sneer

gl-		sn-
glared	glinted	sneered
glaring	glinting	sneering
glaringly	glisten	sneers
glass	glistened	sneezed
glass-bottom	glistening	sneezing
glasses	glitter	snickered
glass-fiber	glittered	sniff
glassless	glittering	sniffed
glass-like	gloom	sniffing
glassy	gloomily	sniggered
glaucoma	gloomy	snippy
glaze	gloss	snivelings
glazed	glossy	snoring
glazes	glow	snorkle
glazing	glowed	snort
gleam	glowered	snorted
gleamed	glowering	snout
gleaming	glowing	snuffboxes
glimmer	glows	snuffed
		snuffer

APPENDIX B

	PRIME	WORD TARGET	NONCE TARGET
PHONAESTHEME	glimmer	gleam	gleat
	glint	glimpse	glim
	glisten	glare	glame
	glitter	glow	glone
	smirk	smell	smett
	smooch	smoke	smofe
	sniffle	sneeze	sneeg
	snore	snack	snab
	snort	sniff	snick
	snout	snarl	snart
	BASELINE	dial	ugly
frill		barn	barm
jest		quill	quive
lure		twin	twim
muck		wrangle	wrongle
poppy		shiver	shigger
sedan		roast	roalt
thirst		hymn	hymp
vigil		yearn	yeart
wry		nest	nisk
MEANING	collar	button	bottin
	cord	rope	rone
	forsake	quit	quin
	keen	clever	cledat
	pollen	blossom	blossot
	somber	dim	diz
	tenor	bass	biss
	thug	bandit	banlit
	varnish	wax	wap
	weed	lawn	lafe

	PRIME	WORD TARGET	NONCE TARGET
FORM	drab	dread	drode
	druid	drip	drit
	flog	flake	flike
	flour	flag	flad
	fray	frost	frossel
	glib	glove	glope
	graze	grunt	grunt
	plight	plunge	plung
	prop	provoke	pronoke
	stout	stab	steb
	PSEUDO-PHON	bleach	blank
blemish		blot	blog
blimp		blob	bloff
bludgeon		blast	blask
blur		blink	bling
crate		cradle	cradon
crony		crook	croof
skipper		skiff	skidge
skit		sketch	skelt
throttle		throat	throam

REFERENCES

- ABELIN, Asa. 1999. Studies in sound symbolism. Göteborg: Göteborg University dissertation.
- BAAYEN, R. HARALD, and ROCHELLE LIEBER. 1991. Productivity and English derivation: A corpus based study. *Linguistics* 29.801–43.
- BALOTA, DAVID; MICHAEL CORTESE; KEITH HUTCHISON; JAMES NEELY; DOUGLAS NELSON; GREG SIMPSON; and REBECCA TREIMAN. 2002. The English lexicon project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. Washington University. Online: <http://lexicon.wustl.edu/>.
- BERKO, JEAN. 1958. The child's acquisition of English morphology. *Word* 14.150–77.
- BLOOMFIELD, LEONARD. 1933. *Language*. New York: Holt.
- BLUST, ROBERT. 1988. Austronesian root theory: An essay on the limits of morphology. Amsterdam: John Benjamins.
- BLUST, ROBERT. 2003. The phonestheme η in Austronesian languages. *Oceanic Linguistics* 42.187–212.
- BOD, RENS; JENNIFER HAY; and STEFANIE JANNEDY (eds.) 2003. *Probabilistic linguistics*. Cambridge, MA: MIT Press.
- BOLINGER, DWIGHT. 1949. The sign is not arbitrary. *Boletín del Instituto Caro y Cuervo* 5.52–62.
- BOLINGER, DWIGHT. 1980. *Language, the loaded weapon: The use and abuse of language today*. New York: Longman.
- BYBEE, JOAN. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- BYBEE, JOAN, and CAROL MODER. 1983. Morphological classes as natural categories. *Language* 59.251–70.
- BYBEE, JOAN, and JOANNE SCHEIBMAN. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37.575–96.
- BYBEE, JOAN, and DAN SLOBIN. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58.265–89.
- CARROLL, LEWIS. 1897. *Alice's adventures in wonderland, and Through the looking-glass*. New York: Macmillan.
- CASSIDY, KIMBERLY; MICHAEL KELLY; and LEE-AT SHARONI. 1999. Inferring gender from name phonology. *Journal of Experimental Psychology: General* 128.362–81.
- CHOMSKY, NOAM. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

- DI SCIULLO, ANNA-MARIA, and EDWIN WILLIAMS. 1987. *On the definition of word*. Cambridge, MA: MIT Press.
- DREWS, ETTA. 1996. Morphological priming. *Language and Cognitive Processes* 11.629–34.
- FELDMAN, LAURIE. 1995. *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum.
- FELDMAN, LAURIE, and EMILY SOLTANO. 1999. Morphological priming: The role of prime duration, semantic transparency, and affix position. *Brain and Language* 68.33–39.
- FIRTH, JOHN. 1930. *Speech*. London: Oxford University Press.
- FRANCIS, W. NELSON, and HENRY KUCERA. 1982. *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- FRISCH, STEPHEN, and BUSHRA ZAWAYDEH. 2001. The psychological reality of OCP-place in Arabic. *Language* 77.91–106.
- HALLE, MORRIS, and ALEC MARANTZ. 1993. Distributed morphology and the pieces of inflection. The view from Building 20, ed. by Ken Hale and Sylvain Bromberger, 111–76. Cambridge, MA: MIT Press.
- HAMANO, SHOKO. 1998. *The sound-symbolic system of Japanese*. Cambridge: Cambridge University Press.
- HAYES, BRUCE. 1980. *A metrical theory of stress rules*. Cambridge, MA: MIT dissertation.
- HOCK, HANS HENRICH, and BRIAN D. JOSEPH. 1996. *Language history, language change, and language relationship: An introduction to historical and comparative linguistics*. Berlin: Mouton de Gruyter.
- HOCKETT, CHARLES. 1954. Two models of grammatical description. *Word* 10.210–31.
- HUTCHINS, SHARON SUZANNE. 1998. *The psychological reality, variability, and compositionality of English phonestemes*. Atlanta: Emory University dissertation.
- JAKOBSON, ROMAN, and LINDA WAUGH. 1979. *The sound shape of language*. Bloomington: Indiana University Press.
- JANDA, RICHARD. 1983. 'Morphemes' aren't something that grows on trees: Morphology as more the phonology than the syntax of words. *Chicago Linguistic Society* 19.2.79–95.
- JANSSEN, THEO. 1997. Compositionality. *Handbook of logic and language*, ed. by Johan van Benthem and Alice ter Meulen, 417–74. Cambridge, MA: MIT Press.
- JOSEPH, BRIAN D. 1983. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general, and historical linguistics*. Cambridge: Cambridge University Press.
- JOSEPH, BRIAN D. 1987. On the use of iconic elements in etymological investigation: Some case studies from Greek. *Diachronica* 4.1–26.
- JOSEPH, BRIAN D. 1994. Modern Greek *ts* beyond sound symbolism. *Sound symbolism*, ed. by Leanne Hinton, Johanna Nichols, and John J. Ohala, 222–36. Cambridge: Cambridge University Press.
- KELLY, MICHAEL. 1988. Phonological biases in grammatical category shifts. *Journal of Memory and Language* 27.343–58.
- KELLY, MICHAEL. 1992. Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review* 99.349–64.
- KELLY, MICHAEL, and KATHRYN BOCK. 1988. Stress in time. *Journal of Experimental Psychology: Human Perception and Performance* 14.389–403.
- KELLY, MICHAEL; KEN SPRINGER; and FRANK KEIL. 1990. The relation between syllable number and visual complexity in the acquisition of word meanings. *Memory and Cognition* 18.528–36.
- KEMPLEY, S. T., and JOHN MORTON. 1982. The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology* 73.441–45.
- LANDAUER, THOMAS; PETER FOLTZ; and DARRELL LAHAM. 1998. Introduction to latent semantic analysis. *Discourse Processes* 25.259–84.
- LANGACKER, RONALD. 1991. *Concept, image, and symbol: The cognitive basis of grammar*. New York: Mouton de Gruyter.
- LIEBER, ROCHELLE. 1980. *On the organization of the lexicon*. Cambridge, MA: MIT dissertation.
- MAGNUS, MARGARET. 2000. *What's in a word? Evidence for phonosemantics*. Trondheim, Norway: University of Trondheim dissertation.
- MARCHAND, HANS. 1959. Phonetic symbolism in English word formations. *Indogermanische Forschungen* 64.146–68.

- MCCUNE, KEITH. 1983. The internal structure of Indonesian roots. Ann Arbor: University of Michigan dissertation.
- MEUNIER, FANNY, and JUAN SEGUI. 1999. Morphological priming effect: The role of surface frequency. *Brain and Language* 68.54–60.
- OHALA, JOHN J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41.1–16.
- PARTEE, BARBARA. 1995. Lexical semantics and compositionality. Invitation to cognitive science, 2nd edn., part 1: Language, ed. by Lila Gleitman and Mark Liberman, 311–60. Cambridge, MA: MIT Press.
- PINKER, STEVEN, and ALAN PRINCE. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28.73–193.
- PLAUT, DAVID, and LAURA GONNERMAN. 2000. Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes* 15.445–85.
- RHODES, RICHARD. 1981. On the semantics of Ojibwa verbs of breaking. Papers of the twelfth Algonquian conference, ed. by William Cowan, 47–56. Ottawa: Carleton University Press.
- RHODES, RICHARD. 1994. Aural images. Sound symbolism, ed. by Leanne Hinton, Johanna Nichols, and John J. Ohala, 276–92. Cambridge: Cambridge University Press.
- RHODES, RICHARD, and JOHN LAWLER. 1981. Athematic metaphors. *Chicago Linguistic Society* 17.318–42.
- SERENO, JOAN. 1994. Phonsyntactics. Sound symbolism, ed. by Leanne Hinton, Johanna Nichols, and John J. Ohala, 263–75. Cambridge: Cambridge University Press.
- SHERMAN, DONALD. 1975. Noun-verb stress alternation: An example of lexical diffusion of sound change in English. *Linguistics* 159.43–71.
- STAROSTA, STANLEY. 2003. Do compounds have internal structure? A seamless analysis. Explorations in seamless morphology, ed. by Rajendra Singh and Stanley Starosta, 116–49. New Delhi: Sage Publications.
- THOMPSON-SCHILL, SHARON; KENNETH KURTZ; and JOHN GABRIELI. 1998. Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language* 38.440–58.
- WALLIS, JOHN. 1699. Grammar of the English language. 5th edn. Oxford: L. Lichfield.
- ZIMMER, KARL. 1969. Psychological correlates of some Turkish morpheme structure conditions. *Language* 45.309–21.
- ZWITSERLOOD, PIENIE. 1996. Form priming. *Language and Cognitive Processes* 11.589–96.

Department of Linguistics
University of Hawai'i at Manoa
569 Moore Hall
1890 East-west Rd.
Honolulu, HI 96822
[bergen@hawaii.edu]

[Received 29 August 2002;
revision received 23 June 2003;
accepted 6 December 2003]