

Phonesthetics and the etymologies of *blood* and *bone*¹

JOSEPH PENTANGELO 

The Graduate Center, CUNY

(Received 25 October 2018; revised 29 October 2019)

The etymologies of English *blood* and *bone* are obscure. Although their cognates are well represented in the Germanic family, both lack clear cognates in other Indo-European languages. Various explanations for their origins have been proposed, including that they may be non-Indo-European (e.g. Hawkins 1987). *Blood* and *bone*, and their cognates, share an initial /b/ with numerous body-related words (e.g. *beard*, *breast*, *bosom*) throughout Germanic. This initial /b/ constitutes a phonestheme. Phonesthemes – ‘recurring sound-meaning pairings that are not clearly contrastive morphemes’ (Bergen 2004: 290) – are present in many Germanic languages, but their role in lexicogenesis is little understood. I suggest that *blood* and *bone* were formed by blending the initial /b/ phonestheme with two pre-existing lexemes: Proto-Germanic **flōda*- ‘something that flows’ and **staina*- ‘stone’. Phonesthetic blending may be a fruitful avenue for future etymological research.

Keywords: Germanic, phonestheme, etymology, word formation, blends

1 Introduction

The etymologies of two common Germanic words, English *blood* and *bone*, are marked by their obscurity. Although their cognates are well represented within the Germanic family, both lack clear cognates in other Indo-European languages. Various explanations for their origins have been put forward, including the claim that neither is Indo-European (Hawkins 1987).

Figure 1 shows the prominent proposed etymologies of *blood*, plus some of its cognates, which are found in all three branches (Eastern, Northern and Western) of Germanic.

As an anatomical term, *bone* refers to elements of the skeleton, but its cognates are polysemous: in Dutch and the Scandinavian languages, it may mean ‘bone’ or ‘leg’, while in German, it predominantly means the latter. (*Knochen* is the regular German word for ‘bone’.) No cognate is found in Gothic. Figure 2 shows its widespread etymologies and some of its cognates.

The association in Old Norse between *beinn* ‘straight’ and *bein* ‘bone, leg’ is dubious; the *Oxford English Dictionary Online* (hereafter *OED*) describes it as possible but unsubstantiated, and calls into question the origin of the Old Norse adjective itself;

¹ Many thanks to Juliette Blevins. Thanks also to Michael Sargent, Bill Haddican, Patrick Honeybone and this article’s two anonymous reviewers, all of whom gave me valuable advice. In 2018, I presented an earlier version of this article at the Word-Formation Theories III conference in Košice, Slovakia. An earlier version of this article won the 2018 Richard M. Hogg Prize from the International Society for the Linguistics of English. Dedicated to my wife, Rebecca.

Orel (2003: 32) and Kroonen (2013) call for considerable semantic elasticity to account for the name of a basic body part, and neither seems especially confident in their disparate accounts.

Gothic *bloþ*, Old English *blōd*, Old Norse *blóð*, Dutch *bloed*, German *Blut*
Proto-Germanic **blōda-* ‘blood’

Derived from Proto-Germanic **blēan-* ‘to blow’, relating to the notion of gushing (*OED*; Skeat 1887; Kroonen 2013), or **blōan-* ‘to flower’, relating to blushing (Kroonen 2013), or **blēda-* ‘breath’, with an association with ‘life’ (Kroonen 2013). Possibly related to Proto-Indo-European (PIE) **b^hleH-* ‘swell, blow up, bubble’ (Boutkan & Siebenga 2005). Watkins (2011) suggests PIE **bhel-* ‘to thrive, bloom’, suffixed form PIE **bhlō-to-* ‘possibly in the meaning’ ‘swell, gush, spurt’. Orel (2003) provides the same analysis morphologically, but glosses **bhel-* as ‘to swell’. Lloyd *et al.* (1998: 211–12) largely agree, but gloss **b^hlō-* as ‘well, gush, swell’. Boutkan & Siebenga (2005) suggest that this word may not descend from PIE at all.

Figure 1. *Blood*’s etymology

Old English *bān*, Old Norse *bein* ‘bone, leg’, Dutch *been* ‘bone, leg’,
German *Bein* ‘leg, bone (archaic)’
Proto-Germanic **baina-* ‘bone, leg’ (Kroonen 2013; but see below)

If connected to Old Norse *beinn* ‘straight’, it may be derived from Proto-Indo-European **bheh₂-* ‘to shine’, assuming that Proto-Germanic **baina-* originally meant ‘beam, ray, post’ (Kroonen 2013). Alternatively, original meaning in Proto-Germanic may have been ‘long bone of the leg’ (*OED*) or just ‘bone’ (Urban 2015; Orel 2003; Lloyd & Springer 1988). May also come from Proto-Indo-European **bhei(ə)-* ‘to strike’ (Orel 2003; Lloyd and Springer 1988). Cate-Silfwerbrand (1958) argued for a Celtic loanword derived from the Proto-Celtic **bend*, **bnd-no-* ‘projecting tip, horn’ (Lloyd & Springer 1988: 516); Hawkins (1987: 75) claims non Indo-European etymology.

Figure 2. *Bone*’s etymology

Both *blood* and *bone* (and their cognates) share an initial *b-* with many other body-related words throughout Germanic (e.g. *beard*, *brain*, *breast*). These words constitute what Dwight Bolinger (1940: 65) dubbed a ‘word constellation’: a group of words sharing similar semantics and a certain phonetic characteristic – in this case, an initial /b/. I suggest that the association of sound and meaning played a critical role in the lexicogenesis of *blood* and *bone*, formed by blending the initial *b-*, suggestive of the group of body-related words, with two pre-existing lexemes: Proto-Germanic **flōda-* ‘something that flows’ and **staina-* ‘stone’, both of which are uncontroversially derived from Proto-Indo European (hereafter PIE) (Boutkan & Siebenga 2005; Watkins 2011: 87; Kroonen 2013).

In section 2, I provide an overview of phonesthesia, with a particular focus on the role of phonesthemes in word formation. In this section, I describe the *b-* ‘body-related’

phonesthetic group. In section 3, I discuss Germanic vocabulary, with consideration given to the Germanic Substrate Hypothesis and the word-formation process of blending. In section 4, I discuss my proposed etymologies for *blood* and *bone*, outlined above, addressing some potential problems with these etymologies. I conclude in section 5 by considering the significance of these suggestions for the consideration of phonesthemes in linguistics more generally.

2 Phonesthesia

Phonesthemes (sometimes *phonaesthemes*) are ‘frequently recurring sound-meaning pairings that are not clearly contrastive morphemes’ (Bergen 2004: 290). The term was coined by Firth in 1930, but the phenomenon it applies to has been described in English since as far back as 1653, when John Wallis included a list of evocative sound clusters in his *Grammatica Linguae Anglicanae*. Some examples on Wallis’ list were *wr-*, showing ‘obliquity or twisting’, as in *wry*, *wrong*, *wreck* and *wrist*, and *br-*, evoking a ‘violent and generally loud splitting apart’, as in *break*, *breach* and *brook* (Magnus 2013: 198). Wallis argued, as Bolinger (1940) would three centuries later, that the meanings of some words could be ascertained through the sound clusters of which they are composed; in *sparkle*, for instance, the *sp-* ‘indicates dispersion’, *-ar-* evokes ‘high-pitched crackling’, *-k-* indicates ‘sudden interruption’ and *-l* ‘frequent repetition’, as in *wiggle*, *wobble* and *twiddle* (Magnus 2013: 199). Most phonesthemes in English are onsets or initial consonants, but rimes and codas may be phonesthetic as well (Firth 1930: 185; Lawler 2006: 1–2). Phonesthemes can occur in any lexical category (Kwon & Round 2015: 14).

The question of where phonesthemes originate remains unanswered. Benczes (2019: 74–83) provides a good overview of the topic. Boussidan *et al.* (2009: 36) suggest that they may have begun as morphemes in a proto-language, which ‘may have survived through generations’. Watkins (2011) lists several Indo-European roots as the progenitors of some Germanic phonesthemes (see section 2.1, below). Blust (2003: 199–200) thoughtfully considers this topic, but concludes that ‘the origin of phonesthemes remains enigmatic’.

There is considerable debate over whether phonesthemes are morphemic (Benczes 2019: 84). Blust (2011: 407) characterizes phonesthemes as ‘submorphemes’, because they ‘can be identified by recurrence, but not by contrast’. Kwon & Round (2015: 24) review this issue, and find that phonesthemes behave like morphemes in most ways, other than the fact that they often appear in ‘lexical stems which are composed of a recurring sound-meaning pairing plus a non-recurrent residue’. Ultimately, they assert that ‘it is imperative that phonaesthemes be accorded a coherent place in morphological theory’. It is sufficient for my purposes here to state that a phonestheme expresses a ‘recognizable semantic association’ without necessarily being classified as a morpheme.

Phonesthemes are often thought of within the domain of sound symbolism, a broad field that also encompasses onomatopoeia and ideophones, defined by the hypothesis that ‘the meaning of a word is partially affected by its sound (or articulation)’ (Magnus

2013: 192). To some extent, sound symbolism is at odds with the notion that the relation between the signifier and the signified is arbitrary, but Blust (2003: 201) sets phonesthemes apart from onomatopoeia, from which they ‘appear to be entirely independent’. Instead, their form is arbitrary: their semantic associations arise from their ‘use and application to new words in the lexicon’, and not some inherent psychological association of sound and meaning (Williams 2013: 597). Firth (1957: 198) railed against associating phonesthemes with ‘the fallacy of sound symbolism’, positing only that ‘a definite correlation can be felt and observed between the use and occurrence of certain sounds and sound-patterns ... and certain characteristic common features of the contexts of experience and situation in which they function’ (1957: 45). There is no empirical evidence that phonesthemes tend to occur in any particular semantic domains (Blust 2003: 201), which further distinguishes them from onomatopoeia, which is relegated to imitative sounds. Nonetheless, many researchers still associate phonesthemes with sound symbolism, going to far as to identify the phonestheme as ‘a type of sound symbolic entity’ (Abelin 2015: 20). Drawing a firm distinction between phonesthemes and sound symbolism is troublesome because sound symbolism is inconsistently defined (Elsen 2017: 491–2).

In order for a phonestheme to exist, there must be a set of words that share similar semantics and a similar phonological form. Bolinger (1940: 65) refers to these groups as ‘word constellations’; the term used in this paper is ‘phonesthetic group’. A phonesthetic group exists regardless of its composite words’ ‘etymology and language of origin’ (Wright 2012: 5), and may contain words belonging to different lexical categories. The number of words necessary to constitute a phonesthetic group is not defined, but the larger the group, the more canonical it is understood to be (Kwon & Round 2015: 13). A phonestheme’s location within a word is important: a given phoneme must appear in a particular position in a series of words with a shared semantic domain, such as the onset of the first syllable or coda of the final syllable, for it to become associated with said domain. In Germanic languages, most phonesthemes occupy the beginnings of words.

Bolinger (1975: 219) writes that the strength of a phonesthetic group can influence the meaning of a word that originally shared with the group a formal, but not semantic, feature. For example, *twiddle*, first attested in 1547, originally meant ‘to be busy with trifles’. However, it acquired the sense of ‘rotate or turn’ around 1676, due to its formal similarities with words like *twist* and *twirl* (Smith 2014: 25). The *OED* is unclear on *twiddle*’s etymology, and suggests that it is actually a blend of *twist* or *twirl* with *fiddle* or *piddle* – however, the fact that it showed no apparent semantic connection to twisting until over a century after its first attestation might call this account into question.

It has also been argued that phonesthesia can be a deciding factor in which words are borrowed. Firth (1930: 191) opined that ‘the importance of “phonaesthemes” in permanently naturalized borrowed words has not been properly recognized’. Carling & Johansson (2014: 211) write that ‘a number of words in sound symbolic [i.e. phonesthetic] networks are loan words’, noting that many cases are inter-Germanic, such as Swedish *glas* ‘glass’ from Middle Low German *glas* ‘glass’ from Old Norse

gler, all in the *gl-* ‘light-related’ phonesthetic group. Others are from outside of Germanic, like English *glair* ‘white of an egg’ and *glairy* ‘wisced, slimy’, from Old French *glaire* ‘egg white’ (ibid.).

2.1 Phonesthemes in English and their role in word formation

Three well-attested English phonesthemes are *gl-*, *sn-* and *gr-*. *gl-* suggests luminousness; it appears in words such as *glisten*, *glow*, *gleam*, *gloss*, *glimmer* and *glitter*. Bolinger (1965: 221–2) estimated that half of the common English words beginning with this cluster had to do with ‘light/vision’. Bergen (2004: 293) consulted an online version of *Webster’s 7th Collegiate Dictionary* and found that 39 percent of word types and 60 percent of word tokens beginning with *gl-* related to ‘light’ or ‘vision’, and that 28 percent of word types and 19 percent of word tokens beginning with *sn-* had definitions relating to ‘nose’ or ‘mouth’. This was described as an ‘overwhelming statistical pairing’. Working with the *Middle English Dictionary* (hereafter *MED*), Williams (2013: 599) found most *gl-* words in Middle English to fall within five main semantic fields: ‘light/vision’ (*glisnen*, *glou*), ‘joy/gladness’ (*gladful*, *glē*), ‘vitreousness/viscosity’ (*glas*, *gleu*); ‘quick/smooth movement’ (*glīden*, *glent*); and deceptiveness (*glōse*, *glāberer*). Tabulating all Middle English *gl-* words in the *MED*, he found the ratio of phonesthetic to questionably/non-phonesthetic words to be 151:84 (i.e. nearly 2:1). Williams also identified certain words as bridging these semantic categories. For example, *glem* connoted both brightness and deception; while it literally meant ‘a beam or radiance of emitted light’, it also indicated ‘a type of what is evanescent or fleeting’, as in the phrase *maken a glem*, ‘to make a deceptive show’ (Williams 2013: 603). Williams found that these polysemous linking words were employed at key points in the Middle English *Pearl* (late fourteenth century) to heighten the poem’s effect.

Sn- suggests an association with the nose, as in *snot*, *snort*, *snout*, *snore* and *sniffle*. According to Philips (2011: 1123), approximately a third of all lexical stems beginning with *sn-* in the *New Short Oxford English Dictionary* have to do with nasality. Francis & Kucera (1982) found 28 percent of word types and 19 percent of word tokens beginning with *sn-* in the Brown corpus to have meanings related to ‘nose’ or ‘mouth’, a percentage far above chance (Bergen 2004: 293). The association of *sn-* with the nose may account for the modern form of *sneeze*: this word is a cognate of the Dutch *fniezen*, Danish *fnyse* and Swedish *fnysa*, ‘to snort’. In Middle English, the word was *fnese*, from the Old English *fnēsan*, ‘to sneeze, puff, snort’; it’s not attested as *sneeze* until 1493. The transformation of Middle English /f/ to English /s/ is not the result of a regular process of sound change; it is accounted for by the semantic pull of other nose-related *sn-* words, like *snore*, *snoke* ‘to snuff or smell’, and *snite* ‘to clean or wipe the nose’ (Burridge & Stebbins 2015: 136).

Gr- suggests ‘grasping’, appearing in such words as *grasp*, *grip*, *grab*, *grapple*, and *grope* (Kwon & Round 2015: 16). Piotr Sadowski (2001) studied *gr-* in Middle English alliterative verse, identifying six main clusters of meaning: hand-object contact

(*graspen, gropen*), ‘the processes of natural life occurring above the ground’ (*gras, ground*), words pertaining to ‘the inside of the earth and things underground’ (*grave, gravel*), agriculture words (*grist, grain*), words involving negative emotions relating to fear (*grendel, grim*), and words involving negative emotions relating to sadness (*greven, groningen*). Like Williams, Sadowski found these groups connected by linking words. He also determined that the majority of phonesthetic *gr-* words were of native Anglo-Saxon origin.

Because these three phonesthemes appear in several other Germanic languages (Blust 2003: 188; Abelin 1999: 135; Firth 1957: 45), it is suggested that their phonesthesia dates back to Proto-Germanic (Carling & Johansson 2014: 206). Watkins (2011: 29, 84) suggests that *gl-* and *sn-* be traced back to PIE, reconstructing their etyma as **ghel-* ‘to shine’ and **snu-* ‘imitative beginning of Germanic words connected to the nose’.

2.1.1 -g ‘animal name’

One of the more interesting English phonesthemes is the final -g in the names of several animals: *dog, frog, pig, stag, earwig, teg* (‘a sheep in its second year’), *hog* (and its compounds, like *warthog*), *bug* and *slug*. This phonesthetic group is presented in table 1, divided into five subgroups. The first two contain names that have been in the group continuously since Old English. Subgroup 3 contains more recent additions. The fourth contains *sucga*, a group-member in Old English without a descendant in the language today. The final subgroup contains *bagga*, a group member in Old English whose modern form, *badger*, does not end in -g, and which is therefore not a member of this phonesthetic group any longer. Some of the Old English forms, marked with an asterisk, have been reconstructed from placenames (Hogg 1982: 195).

Table 1. *The -g animal names*

English	Old English
1 <i>dog, frog, pig, stag, earwig, teg</i>	<i>docga, frogga, *picga, *stacga, (ēar-)wicga, *tacga~tecga</i>
2 <i>hog</i> (+ its compounds, e.g. <i>hedgehog</i>)	<i>hogg~*hogga</i>
3 <i>bug, slug</i>	N/A
4 N/A	<i>sucga</i> ‘hedge sparrow’
5 N/A [<i>badger</i>]	<i>bagga</i> ‘badger’

2.1.2 Old English -cga, -gga ‘hypocoristic animal name’

According to the *OED*, the words in subgroup 1 form a set ‘of uncertain or phonologically problematic etymology’ dating back to Old English, where each of them contained the medial [gg] geminate, written as -cg- or -gg-. Hogg (1982: 195) highlights this geminate’s rarity: including actual samples plus those reconstructed from place names,

it appeared in only 21 Old English words. The majority of geminates in Old English are attributed to West Germanic doubling of a consonant before *j; however, a [gg] cluster in this environment would have subsequently been palatalized early on in the development of Old English. The fact that these words contain [gg], as opposed to [dʒ], indicates that they cannot have resulted from West Germanic gemination, meaning that their development was the result of an innovative process taking place within Old English itself.

A significant portion of all Old English [gg] words are animal names that take the form of masculine weak nouns, including the etyma of the words listed above: *docga*, *frogga*, **picga*, **stacga*, (*ēar-*)*wicga* and **tacga* ~ **tecga*, as well as *sucga* ‘hedge sparrow’.² Gąsiorowski (2006: 279) suggests that *bagga* ‘badger’ be added to this list, and the *OED* would add *hogga* ‘hog’, a strong masculine that may originally have been weak. Most of these words have etymologically transparent, ‘more important’ synonyms in Old English: *hund*, *frosc* (‘the normal form in the Germanic languages’) or *frox*, *swīn*, *heorot* ‘male deer’, *ceafer* ‘beetle, locust, caterpillar, or other pest’, *scēap* or *ēowu* and *spearwe* ‘sparrow’, which they only came to replace gradually (Hogg 1982: 196). As a result, there is general consensus that the [gg] animal names were originally hypocoristic forms (Gąsiorowski 2006: 280; Hogg 1982: 196; *OED*). It is worth mentioning here that the medial [gg] is not apparently imitative in any way. Rather, it may be understood as a phonestheme, one of several ‘frequent pairings of phonemes and aspects of meanings’ (Elsen 2017: 492), in this case expressing the fact that a given animal name is hypocoristic.

2.1.3 *Development of the phonesthemic group*

Slug has perhaps the simplest etymology, and appeared the latest. From the Middle English adjective *slugge* ‘to be lazy, slow, or inert’, a likely Scandinavian borrowing, it is not attested as an animal name until 1703 (*OED*). Although its derivation has nothing to do with its phonesthemic quality, this is a member of the group today due to its shared semantic and phonological affinities.

While *hog* is generally considered a Celtic borrowing (see Welsh *hwch*, Cornish *hogh*), it is suggested by both Hogg (1982: 197) and the *OED* that it ‘may have been partially assimilated to the group of which **picga* is a member, on purely semantic grounds’. This semantic pull may also be responsible for *frogga*, derived by replacing the final consonants of the original *frosc*~*frox* ‘frog’ with the -*gga* phonestheme (*OED*; Gąsiorowski 2006: 280). *Bug* ‘insect’, first attested in 1622, was influenced by this process as well. The *OED* describes its etymology as ‘origin unknown’, and notes:

Perhaps a transferred sense of *bug* [‘an imaginary evil spirit or creature’], insects being taken to resemble typical representations of monsters or the monstrous, although early examples show no clear evidence of such association. Sense 1 [‘any small insect or larva that is considered to be a pest’] also appears to show either connection or confusion with earlier *budde* and *boud*, which occur in similar senses: beside *sharn-bug* compare earlier *sharnbud*.

² *Wicga* survives in *earwig*, but was a general term for an insect or beetle.

Just as *frosc*~*frox* became *frogga* within Old English, earlier *bud* became English *bug* – and both transformations were apparently motivated by the words’ semantic connections with the same phonesthetic group.

Brian D. Joseph (1997: 12) calls this phenomenon ‘phonesthematic attraction’, applicable when ‘sound symbolic clusters of words ... draw other words into their “orbit”, so that these other words change their form in the direction of the sound symbol’. More poetically, Bolinger (1953: 328) describes this as ‘a change of form to make the word seem to mean what it really means’.

Using the analogical transformation of *frosc*~*frox* to *frogga* as a starting point, Gašiorowski (2006: 281–2) suggests that *docga* is derived from *dox*~*dohx* ‘yellowish-brown’, an appropriate source given the fawn or brindle color of the mastiff. The usage of color words for hypocoristics is well attested in Old, Middle and Modern English: Gašiorowski (2006: 280) mentions *Blæcca*, from *blæc*, an Old English nickname for someone with black hair; *Bruin* ‘bear’, printed by Caxton in *Reynard the Fox* (1481), relates to *brown*; and *Red*, *Blondie* and *Blackie* are commonly used to refer to people or animals with hair of those colors today.

Under Gašiorowski’s analysis, the lexeme *dox* ‘yellowish-brown’ was taken as the base for the creation of a new term for an animal of that color. The new lexeme was intended to fit into a group of other animal names, all of which ended in *-cga* or *-gga*. Its final consonant sounds were replaced with *-cga*, thereby making it conform to the rest of the series (2006: 281–2). Gašiorowski analyses this derivation as a strengthening of the final obstruent in *dox*, /ɣ/ or /x/, arising from a hypocoristic truncation of the underlying word. While this strengthening can account for the forms of *docga* and *frogga*, this does not explain the many other animal names in this phonesthetic group. I suggest instead that the base was blended with the *-cga* phonestheme (see figure 3).

$$\begin{array}{rcccl} \textit{dox} & & \textit{x} & & \textit{-cga} & = & \textit{docga} \\ \text{‘yellowish-brown’} & & & & \text{‘hypocoristic} & & \text{‘dog’} \\ & & & & \text{animal name’} & & \end{array}$$

Figure 3. *Dog*

This process is similar to that by which *frosc* yielded *frogga*, and by which *fnese* became *sneeze*, but differs in a very important way. *Frosc* always meant ‘frog’ and *fnese* always meant ‘sneeze’ – their phonological forms were changed in order to fit them into phonesthetic groups, but their meanings remained the same. But with *docga*, the resulting word’s meaning, ‘dog’, is completely different from that of *dox* ‘yellow-brown’. Although basically synonymous with the pre-existing *hund*, it constitutes an entirely new word, given its completely distinct phonological form and its difference in register.³

³ Similarly, *puma*, *cougar*, *catamount* and *mountain lion* are four different names for the exact same animal, *Puma concolor*.

2.2 Analogical emergence

Phonesthemes may also be used in the creation of new words through the process of analogical emergence. Carling & Johansson (2014: 210) named this method, wherein ‘words are created by means of an association to other sound symbolic words within the language, formally and semantically’. In practice, a speaker incorporates a phonestheme in the creation of a new word. The inclusion of the phonestheme helps to express something of the word’s intended meaning, provided that the listener is familiar with the corresponding phonesthetic group.

Åsa Abelin (1999: 232–7) has conducted experiments wherein subjects were tasked with coining nonce words carrying a particular meaning. Semantically appropriate phonesthemes were usually incorporated into the subjects’ creations. In other words, phonesthetic patterns are accessible to speakers, and are used productively in word formation, even though speakers might not be consciously aware of their existence. Later, Abelin (2015: 24–7) conducted an experiment wherein ten Swedish-speaking subjects were tasked with pairing nonce words with pictures. The pictures were each chosen to reflect a trait associated with a given phonestheme; this phonestheme was present in one of the two nonce words made available for each picture. Subjects usually matched the images with the words that contained the semantically appropriate phonestheme – for example, when presented with an image of a knotty lump of wood, eight of the ten participants chose the word *skrob* over *blik*, in correspondence with the Swedish *skr-* ‘rough surface’ phonestheme (Abelin 2015: 27).

Because they grant listeners an immediate understanding of a previously unknown word, phonesthemes are popular among professionals who create names for products and companies. Abelin (2015: 26–7) conducted studies using the corpora of the Swedish Patent and Registration Office and the Swedish registry of medicines, with the database of written Swedish (KORP) used to establish a baseline. It was found that pejorative phonesthemes are almost uniformly avoided in Swedish brand names, and that semantically appropriate phonesthemes were overrepresented in certain fields: for example, *fl-* ‘speed-related,’ appeared often in names for medications in Sweden.

But incorporating phonesthemes in the formation of new words is not limited to nonce words and brand names, as the English *-ash* ‘forceful strike’ phonesthetic group illustrates.

The words in this robust group were introduced chiefly in Middle and Early Modern English. Words include *bash*, *clash*, *crash*, *dash*, *gnash*, *hash*, *lash*, *mash*, *pash*, *slash*, *smash* and *swash*. Some of these are only dialectal today, while others are rather high-frequency words. It seems that only *mash* originated in Old English. *Hash* is a borrowing of French *hacher* ‘to cut into small pieces’. Two, *lash* and *slash*, may have been borrowed from Old French (*lascher* ‘to loose, let go’ and *esclachier* ‘to break’ respectively). Others may have Scandinavian origins: the *OED* compares *bash* to Swedish *basa* ‘baste, whip’ and Danish *baske* ‘beat, strike’; *dash* to Swedish *daska* ‘drub’ and Danish *daske* ‘to beat’; and *gnash* is derived from the older English *gnast*, a borrowing derived from Old Norse **gneista* ‘gnash teeth’. But the *OED* cautions that

these cross-linguistic similarities may simply be a case of ‘analogous formations’, noting that English *crash* is similar, but probably not actually related to, Swedish *krasa*, Danish *krase* ‘to crackle’.

The etymologies for most of these words – even those described as possible borrowings – are chiefly characterized by the *OED* as ‘onomatopoeic’. For some of these words, this seems appropriate: *pash* and *smash* are ‘probably imitative’, and *swash* is ‘echoic’ – onomatopoeic in the true sense of the word. But many of these words are not actually imitative, and, bearing in mind that many of these etymologies were written before Firth coined the word ‘phonestheme’, I suggest a change in terminology: rather than onomatopoeia, several of these words were formed through phonesthesia.

For instance, the *OED* suggests that *clash* was formed ‘from instinctive association with classes of pre-existing echoic words. The initial element is that of *clap*, *clack*, etc.; the final that of *dash*, *splash*, *smash*, *swash*, etc., or perhaps a direct imitation of the element of sound common to these.’ These ‘classes of pre-existing echoic words’ might otherwise be called phonesthetic groups. This etymology suggests a compounding of the *cl-* ‘noisy collision’ phonestheme with the *-ash* phonestheme. Similarly, *crash* is etymologized as ‘having the same relation to *crack* that *clash* has to *clack* and *clap*’, in other words taking the *cr-* ‘noisy impact’ phonestheme identified by Bloomfield (1933: 245) as its first element. *Bash* is described as combining ‘the *b* of *beat*, *bang*, and the termination of *dash*, *gash*, *gnash*, *hash*, *lash*, *pash*, *smash*, etc.’. In these three examples, phonesthemes are taken as bases for the formation of entirely new words.

2.3 b- ‘body-related’

I suggest that *b-* is a phonestheme in English, indicating ‘part of the body, body-related’. Magnus (1998) proposed the phonesthetic group of *b-* ‘body parts’, and calculated the words in this group as constituting 5.31 percent of the 583 English words with an initial *b-*.⁴ Notably, she populated this phonesthetic group with fewer words than I do. Several of these words can be seen illustrated in figure 4.

The words that I have placed in the *b-* phonesthetic group, listed in table 2, all feature an initial *b-* and are spread across four primary semantic subdomains relating to the body: external body-parts, internal body-parts (including fluids and gases), bodily injuries or malformations and bodily verbs, as well as *body* itself. A few are body-related adjectives. I also included two words which survive only in dialectal English: *bree* ‘the eyelid’, described in the *OED* as obsolete except in northern dialects, and *bouk* ‘the trunk of the body’, now ‘Scottish and dialectal’. It is also noteworthy that a number of words existed in Old English which no longer make up part of Modern English vocabulary, and are thus absent from table 2. These include kennings like *bānhūs* ‘body [bone-house]’, Latin glosses like *burse* ‘scrotum’, and everyday words like *bæcþearmas* ‘intestines’ and *bearm* ‘bosom, lap’ (Cameron *et al.* 2018; Hough & Kay 2017).

⁴ Magnus does not indicate the source for the 583 English *b-*initial words.



Figure 4. This illustration of the bearded woman of Limerick, from a manuscript containing Gerald of Wales' *Topographia Hiberniae* (c. 1196–1223), displays numerous body-related *b*- words, including *bare*, *body*, *brow*, *beard*, *breasts*, *belly* and *buttocks*. (British Library MS Royal 13 B VIII, f. 19r. British Library Catalogue of Illuminated Manuscripts www.bl.uk/catalogues/illuminatedmanuscripts/welcome.htm. Image is in the public domain.)

The first column lists basic words, while the second lists relevant derivatives, e.g. dialectal variants, clippings, or compounds. The third column provides a brief etymology of each word. For development from Old English, my sources were Hall's (1960) *Concise Anglo-Saxon Dictionary*, Skeat's (1887) *Principles of English Etymology*, vol. I, and the *OED*. Proto-Germanic reconstructions are from Kroonen's (2013) *Etymological Dictionary of Proto-Germanic*. Where Kroonen provided no etymology, I used Orel's (2003) *Handbook of Germanic Etymology*; where Orel was lacking, I turned to Fick *et al.*'s (1909) *Wortschatz der Germanischen Spracheinheit* (cited as 'Fick' in table 2). Although Kroonen's reconstructions take primacy, I consulted all three of these dictionaries for background information and alternative analyses. PIE reconstructions are from Watkins' (2011) *American Heritage Dictionary of Indo-European Roots*, with Fortson's (2010) *Indo-European Language and Culture* consulted as well.

Notes on the reliability of the Proto-Germanic and PIE roots are informed by the sources listed above plus Boutkan & Siebenga's (2005) *Old Frisian Etymological Dictionary* (cited in table 2 as 'Boutkan') and Liberman's (2008) *Analytical Dictionary of English Etymology*.

Language abbreviations include, in order of appearance, OE (Old English), PGmc. (Proto-Germanic), PIE (Proto-Indo-European), ME (Middle English) and IE (Indo-European).

Table 2. *Body-related b-words in English*

Basic word	Derived words	Brief etymology
back	backbone, backside	OE <i>bæc</i> , PGmc. <i>*baka-</i> (Orel). No likely PIE root. Possibly related to Slavic <i>*bokъ</i> 'side'.
ball	bollock (early OE; <i>ball</i> + <i>-ock</i> 'diminutive')	OE <i>beallucas</i> , PGmc. <i>*ballan-</i> 'ball', PIE <i>*bhel-</i> 'to blow, swell; with derivatives referring to various round objects and to the notion of tumescent masculinity'. <i>Bollock</i> is attested before <i>ball</i> , although <i>bollock</i> contains the diminutive <i>-ock</i> suffix, suggesting <i>ball</i> 's precedence. See <i>buttock</i> .
bare beard		OE <i>bær</i> , PGmc. <i>*baza-</i> , PIE <i>*bhoso-</i> 'naked' OE <i>beard</i> , PGmc. <i>*barzda-</i> 'beard'. Strong doubt over origins. Although attested in other IE languages, likely borrowed into Balto-Slavic, and possibly Latin, from Germanic (Boutkan, Kroonen). Watkins posits PIE <i>*bhardh-ā-</i> 'beard'. No consensus.
belch		OE <i>bealcan</i> , <i>bealcettan</i> , <i>bælcan</i> . Unknown etymology. Perhaps related to PGmc. <i>*bulgjan-~*bulkjan-</i> 'to bellow', from PIE <i>*bhel-</i> 'to cry out, yell', but this is debated (Kroonen). Fick gives PGmc. <i>*bel-</i> 'sound, roaring' as the source of OE <i>bealcan</i> , Dutch <i>balken</i> 'screaming of a donkey', <i>bulken</i> 'roar', Middle Dutch <i>bulghen</i> 'burp', <i>bulsen</i> 'cough', ME <i>belsen</i> 'yell', and others, listing PGmc. <i>*buljan-</i> (i.e. Kroonen's <i>*bulgjan</i>) as a derived form.
bell-end belly	belly button	Compound, 1827 OE <i>bælg</i> , <i>belg</i> 'a bag, skin (for holding things)', PGmc. <i>*balgi-</i> 'skin bag', PIE <i>*bhelgh-</i> 'to swell'
bile bite (<i>n.</i> & <i>v.</i>)		French <i>bile</i> , Latin <i>bīlis</i> OE <i>bite</i> , PGmc. <i>*bītan-</i> 'to bite, be sharp', PIE <i>*bheid-</i> 'to split'
bladder		OE <i>blædre</i> 'blister, bladder', PGmc. <i>*blaðron</i> 'bladder' (Orel), PIE <i>*bhlē-</i> 'to blow'
blain		OE <i>blegen</i> . Unknown etymology. <i>OED</i> suggests PGmc. <i>*bleganā-</i> . Not likely IE.

(Continued)

Table 2. (continued)

Basic word	Derived words	Brief etymology
blemish		ME <i>blemyss</i> , <i>blemiss</i> , and other forms, Old French <i>blemiss</i> ‘to render livid or pale’. Further etymology unclear. Watkins suggests Old French borrowed from PGmc. <i>*blas-</i> ‘shining white’, ultimately from PIE <i>*bhel-</i> ‘to shine, flash, burn’
blind		OE <i>blind</i> , PGmc. <i>*blinda-</i> , PIE <i>*bhel-</i> ‘to shine, flash, burn’
blink		ME <i>blynke</i> , ‘occasional variant of [ME] <i>blenk</i> ’, itself from OE <i>blencan</i> ‘to deceive, cheat’, PGmc. <i>*blanka</i> ‘colorless?’ (question mark Kroonen’s), PIE <i>*bhel-</i> ‘to shine, flash, burn’
blister		ME <i>blester</i> , <i>blister</i> , possibly from Old French <i>blēstre</i> ‘tumor’. <i>OED</i> : ‘An Old English <i>blēster</i> , <i>blēster</i> or <i>blýster</i> , cognate with the Old Norse [<i>blástr</i> , <i>blástri</i> ‘swelling’] or Dutch [<i>bluyster</i> ‘blister’], might have been expected, but is not found’. Kroonen (2013) suggests ME borrowed an Old French word which developed from a Latin borrowing from Germanic, <i>*bulgjon</i> .
blood	bleed	OE <i>blōd</i> , PGmc. <i>*blōda-</i> ‘blood’. No clear further etymology – possibly from PIE <i>*bhlō-to-</i> ‘swell, gush, spurt’, derived from <i>*bhel-</i> ‘to thrive, bloom’; or meaning ‘life’, derived from PIE <i>*bhlē-</i> ‘to blow’. See discussion in Kroonen. Widely suggested as non-Indo-European, as in Boutkan.
blow body		OE <i>blāwan</i> , PGmc. <i>*blēan-</i> , PIE <i>*bhlē-</i> OE <i>bodig</i> ; cognates in Old High German <i>potach</i> ‘body, trunk, corpse’, regional German (Austrian, Swabian, Bavarian) <i>Bottig</i> . No proposed PGmc. or PIE etymology.
boil <i>n.</i>		OE <i>býl</i> , PGmc. <i>*bùljō(n)</i> (Orel), PIE <i>*bheuə-</i> ‘to be, exist, grow’ or <i>*bhelgh-</i> ‘to swell’. Unclear etymology; generally considered IE.
bone	boner (1962)	OE <i>bān</i> , PGmc. <i>*baina-</i> ‘bone, leg’. Dubious IE etymology; maybe meant ‘beam, post, ray’, from PIE <i>*bheh₂-</i> ‘to shine’; also possibly from PIE <i>*bheiə-</i> ‘to strike’; see definitions in Kroonen and Orel. Often suggested as non-Indo-European, as in Hawkins (1987: 75).

(Continued)

Table 2. (continued)

Basic word	Derived words	Brief etymology
bosom		OE <i>bōsm</i> , PGmc. <i>*bosmaz</i> ‘bosom, breast’. Only in West Germanic. Not likely IE.
bottom	botty (1874), booty (1926; from botty), batty (1935; from botty, Caribbean)	OE <i>botm</i> , PGmc. <i>*budman</i> ~ <i>*buttman</i> , PIE <i>*bhudh-</i> ‘bottom, base’. <i>OED</i> : <i>bottom</i> as anatomical term dates to approximately 1550.
bowel		ME <i>buel</i> , <i>bouel</i> , Old French <i>boel</i> , <i>buel</i> , <i>bouel</i> , late Latin <i>botellus</i> ‘pudding, sausage, a small intestine’
bouk (‘Now only Scottish and dialectal’)		OE <i>būc</i> , PGmc. <i>*būkaz</i> ‘belly’ (Orel). No likely PIE root. Maybe from <i>*beu-</i> / <i>*bheu-</i> ‘imitative root ... associated with the notion “to swell”’ (Watkins).
brain		OE <i>brægen</i> , PGmc. <i>*bragna</i> , unclear PIE etymology. PIE <i>*mregh-m(n)o-</i> ‘brain’ is the classic etymon, but this has been broadly challenged for phonological reasons. Maybe from PIE <i>*bheragh</i> ‘high’ (Orel), maybe from PIE <i>*bhragno</i> ‘something broken’ (Lieberman 2008), maybe not PIE at all (Boutkan). If not from PIE, then no cognates outside of Western Germanic.
breast		OE <i>brēost</i> , PGmc. <i>*breusta-</i> , <i>*brust-</i> ‘breast, chest’, PIE <i>*bhreus-</i> ‘to swell’
breath	breathe	OE <i>brāþ</i> , <i>brēþ</i> , PGmc. <i>*brēan</i> ‘to fume, smell’, PIE <i>*gwhrē-</i> ‘to smell, breathe’
bree (‘Northern English’)		OE <i>brēw</i> , <i>brēaw</i> , PGmc. <i>*brēwō-</i> , PIE <i>*h₃b^hrēuH-o</i> (Boutkan, Kroonen)
brow		OE <i>brū</i> , PGmc. <i>*brū-</i> ‘bridge’, PIE <i>*bhrū-</i> ‘eyebrow’. Etymology sometimes confused with that of <i>bree</i> ; possibly both are from the paradigm of a shared etymon in PIE (Kroonen).
bruise		OE <i>brýsan</i> ‘to crush, bruise’. Unknown etymology. Watkins provided PGmc. <i>*brūsjan</i> , PIE <i>*bhreu-</i> ‘to cut, break up’
bubby	boob (1908; shortening of bubby; as ‘breasts’, 1949), booby (1934; from bubby)	1690, unknown etymology. Compared with German <i>bübbi</i> ‘teat’.
bum		ME <i>bom</i> . <i>OED</i> : ‘Perhaps of imitative origin; compare other words of similar sound and with the general sense of “rounded protuberance, swelling”, as e.g. <i>bump</i> , <i>bumb</i> [‘a pimple’], <i>bub</i> [‘a small pustule or nodule

(Continued)

Table 2. (continued)

Basic word	Derived words	Brief etymology
bump <i>n.</i>		in or beneath the skin’]. Clearly the <i>OED</i> is here referring to a phonesthetic group. <i>OED</i> : ‘Probably of imitative origin ... compare other words of similar sound denoting a rounded protuberance or swelling as e.g. <i>bum</i> , <i>bumb</i> [‘a pimple’], <i>bub</i> [‘a small pustule or nodule in or beneath the skin’], <i>lump</i> , etc.’. Dated to 1533. No certain etymology.
bunion		First attested 1718. Perhaps from Italian <i>bugnone</i> ‘a push, a bile, a blane, a blotch’ (Skeat), but <i>OED</i> is very skeptical of this. No certain etymology.
burp		<i>OED</i> : ‘imitative’. Dated to 1932. Kroonen: ‘of sound-symbolic origin’, akin to, but not descended from PGmc. <i>*rup(p)ōn-</i> ‘to belch’
bust	busty	First used for sculptures of torsos approximately 1660, as ‘bosom’ 1807. French <i>buste</i> ‘upper part of the trunk’, Italian <i>busto</i> ‘upper part of the human trunk, from the neck to the hips’, possibly from Latin <i>bustum</i> ‘funeral pyre, tomb’
buxom		ME <i>buhsum</i> , probably OE <i>*(ge)būhsum</i> , possibly PGmc. <i>*beugan-</i> ~ <i>*būgan-</i> ‘to bow, bend’. Acquired its modern meaning by 1589.
butt	buttock (1300; <i>butt</i> + <i>-ock</i> ‘diminutive’)	ME <i>buttok</i> . Possibly PGmc. <i>*buttaz</i> ~ <i>*būtaž</i> (Orel), the etymon of Norwegian <i>bútr</i> ‘log’, Low German <i>butt</i> ‘blunt, plump’, and Middle High German <i>butze</i> ‘cut out piece’ (Orel). Possibly related to <i>*bautanan</i> ‘to beat’, and thence PIE <i>*bhau-</i> ‘to strike’ (Orel, Watkins), but this semantic leap is never explained. <i>Buttock</i> is attested before <i>butt</i> , although <i>buttock</i> contains the diminutive <i>-ock</i> suffix, suggesting <i>butt</i> ’s precedence. See <i>bollock</i> .

Nearly all of the words date back to Proto-Germanic, suggesting that this phonesthetic group existed in that language as well. This is somewhat supported by the phonesthetic group’s presence in other modern Germanic languages. For example, German includes many cognates of English, such as *Bälle* ‘balls’, *Bart* ‘beard’ and *Busen* ‘bosom’, but also *Beule* ‘bump’ and *Bauch* ‘stomach’. English does have more words in the *b-* ‘body-related’ phonesthetic group than does German, but this is not exceptional. A comparison can be made to the *sn-* ‘nasal/oral area’

phonesthetic group, which Blust (2003: 188) reckons as containing 19 words in English, 11 in German and 12 in Dutch.

It has been argued that poetry, both alliterative and rhyming, may serve to teach and reinforce phonesthetic associations (Benczes 2019: 76–7). Alliteration is the major characteristic of the oldest surviving Germanic poetry (Lehmann 1971: 4; Fortson 2010: 350). It is attested in Old High German, Old Icelandic, Old Saxon and Old English (Lehmann 1971: 23). Alliteration is found in several of the oldest runic inscriptions, such as that on the fifth-century Gallehus horn: *Ek HlewagastiR HoltijaR horna tawidō* ‘I Hlegestr of Holt made the horn’ (Lehmann 1971: 28). The Ström whetstone from Norway, carved in the early seventh century, includes the inscription *wate hali hino horna hahaska þi haþu ligi* ‘Let the horn moisten this hanging stone, so that the grass may lie’ (Owen 1928: 3). These runes are believed to be the closest surviving approximants of Proto-Germanic (Lehmann 1971: 77). Early runic inscriptions like these date to ‘some stage of development between a relatively homogenous [North-West Germanic]’ and the earliest manuscripts in the differentiated Germanic languages (Findell 2012: 3).

Many *b*- ‘body-related’ words appear together in verse. This is demonstrated in the Old English poem *St Guthlac A* (late tenth century). When demons attack Guthlac, an angel commands them not to harm him: *Ne sy him banes bryce, ne blodig wund* ‘let there be in him no break of a bone, nor bloody wound’ (Gollancz 1895: 147).

The common alliterative grouping of words from this phonesthetic group may have contributed to the coining or borrowing of some *b*-initial body-related words. The *OED* cites the *Cursor Mundi* (c.1325) as the first attestation of *blester* ‘blister’ in Middle English, where it appears alongside *bile* ‘boil’ and *bolnand* ‘swelling (up)’ in a passage about the plagues of Egypt affecting the *bodis* of the pharaoh’s people: *Bile and blester, bolnand sare* ‘boil and blister, swelling sore’.

In England, rhyming largely replaced alliteration after the Norman conquest, although it survived in the north and the west. There was something of an alliteration revival from the thirteenth through sixteenth centuries, perhaps suggesting a continued tradition or Scandinavian influence (Lehmann 1971: 23–4).

3 Germanic vocabulary

From the inception of Indo-European philology, Germanic has been regarded as something of an outsider. Sir William Jones (1798: 423) described ‘the Gothick’ and Celtic languages as ‘blended with a very different idiom’. One of the hallmarks of its apparent otherness is its vocabulary: it is estimated by some scholars that up to a full third of the Germanic lexicon is of non-Indo-European origin (Feist 1914: 88; Hawkins 1987: 71; Kroonen 2011: 126). Hawkins (1987: 74) asserts that these words ‘belong to the very core of the basic vocabulary of Common Germanic’. Others have arrived at far more conservative estimates, including Prokosch (1939: 23), who posits ‘a negligible quantity’ of substratum words, and Kroonen (2013), who cites 15 percent as ‘etymologically unclear’ and only 4–5 percent as explicitly non-Indo-European.

3.1 Germanic Substrate Hypothesis

The Germanic Substrate Hypothesis is a popular but controversial theory that addresses the origins of these etymologically difficult words. It posits that they are the remnants of an extinct substrate language spoken by the natives of northern Europe (Sörög 2015: 13). The Germanic languages, it is argued, retained a large share of words from this substrate, but relics may be found in the Celtic, Slavic, Italic and Baltic families as well (Boutkan 1998: 102). The hypothesis is that ‘the Germanic family emerged from a contact language spoken by both the Indo-European newcomers and indigenous inhabitants’ (Pereltsvaig & Lewis 2015: 138).⁵

Many have argued that this explanation has been grossly over-applied (Roberge 2010; Kroonen 2011: 126–32; 2013; Pereltsvaig & Lewis 2015: 138). Kroonen (2012: 255) supports a conservative version, wherein the Indo-Europeans borrowed agricultural terms from the Neolithic Europeans among whom they settled. Others contend that the theory has essentially no merit at all (Schuhmann 2012).⁶

3.2 Germanic word formation

If a given word has unclear origins, advocates of the Germanic Substrate Hypothesis often suggest that it is a relic of the Pre-Germanic substrate. This does not take into account the fact that new words are frequently coined in living languages, often without recoverable etymologies. In his review of Boutkan & Siebenga (2005), Anatoly Liberman (2006: 4) takes issue with the over-attribution of difficult words to a substrate origin, writing:

In American slang, a state of nervous excitement can be called *tizzy*, *dither*, and *swivet*. Their phonetic shape is somewhat unusual, their related forms have not been found, and their origin, except possibly for *tizzy*, is ‘unknown’. To complicate matters, *tizzy*, recorded only in the 19th century, first surfaced in texts with the meaning ‘sixpence’ (the same word?). *Tizzy*, *dither*, and *swivet* are not substrate words, are they?

Throughout the Germanic languages, the primary methods of lexicogenesis are compounding, derivation (the application of affixes to roots) and borrowing. Affixes may have their origins as independent roots, such as the *-hood* of *childhood*, from Old English *hād* ‘person, personality, sex, condition, quality, rank’ – as a result, synchronic

⁵ Sigmund Feist (1932) is often credited as the theory’s originator. Although he asserts that ‘the Pre-Germans ... had previously spoken a different language’ (1932: 248) than the Indo-Europeans, he makes no claims regarding a special status of Germanic or of a particularly noteworthy linguistic substrate. His theory, in his own words, is ‘that to the Pre-Germans of northern Europe speech as well as writing was brought by the Veneti-Ilyrii’ (1932: 251).

⁶ Along similar lines, Theo Vennemann argues that a Vasconic substrate accounts for many of the words found chiefly in Celtic, Germanic and Italic, but not as much in other Indo-European families (Vennemann 2003: 343–4). Furthermore, he argues that a Semitic superstrate (probably Punic), introduced by seafaring Phoenicians, is responsible for many of the idiosyncratic words in Germanic vocabulary (Vennemann 2012: 436). The argument advanced against the Germanic Substrate Hypothesis by Liberman (2006), below – that novel words may arise from internal developments – has also been brought against Vennemann’s analyses (Baldi & Page 2006: 2190).

derived forms may have been compounds when they were first created. New words may also be introduced by several relatively minor processes; relevant to this article are blending and phonesthesia.

3.2.1 *Blending and phonesthesia*

Blending is an inexact process whereby at least two elements are combined to create a single, new lexeme, known as either a 'blend' or 'portmanteau' (Bauer 2006: 502). Bat-El (2013: 371) notes that 'blends are somewhat like compounds, but with fewer restrictions'. Blending is found in many languages, including English, Russian, Icelandic, German and Hebrew (Tappenden 2009; Pereltsvaig 2010; Bat-El 2013).

Most blends use two separate lexemes as their elements. Typically, these elements are clipped word-internally at the blend's 'switchpoint', usually a place of phonetic or graphemic overlap (Gries 2004: 645). For example, the switchpoint of *spork* is *o*, found in both elements, *spoon* and *fork*. Blends, especially those without overlapping segments, usually bear the prosody of the longer of the two elements. If a polysyllabic blend's first element is monosyllabic, it will not usually be clipped (e.g. *foolosopher*). If its first element is polysyllabic, but can be fit into the prosodic structure of the blend's second element, it usually won't be clipped either (e.g. *dramedy*). Blends are frequently used in the media, as product names, and as scientific and technical terms (Szymanek 2005: 434).

Chris Smith (2014) explored the role of phonesthesia in blends, and found that 55 percent of blends coined between 1200 and 1900 fit within phonesthetic groups. For example, eight blends fit within the *fl*- 'motion, repeated or fluid' phonesthetic group: *flaunt*, *flounder* v., *furry*, *flush*, *flare*, *flustrate*, *fluff* and *flimmer*. This is unsurprising: if the first element contains an initial phonestheme, or if the second element contains a final phonestheme, then the resulting blend should contain that phonestheme too. More interesting is Smith's finding that blends are often reanalyzed to fit into phonesthetic groups that their elements might not have belonged to, especially when the blend's form is opaque enough that its elements are hard to recognize. In other words, phonesthematic attraction commonly asserts itself on blends. Another interesting finding is that only 1.5 percent of the 202 blends coined after 1900 seemed to be phonesthetic. This is attributed to the more recent blends tending to be more transparent, and thus less likely to be reanalyzed (Smith 2014: 29). It is also possible that blends belonging to phonesthetic groups tend to be longer lasting, perhaps owing to their phonesthesia. Smith uses the *OED* as the source for pre-1900 blends, but contemporary research for those coined after 1900. As a result, there is an imbalance in the blends studied: prejudice in favor of well-attested, long-lasting pre-1900 words is mixed with a *laissez faire* acceptance of more recent neologisms, regardless of their popularity and longevity.

As discussed in section 2.1.3, blends may be composed of a phonestheme and a lexeme from the outset. Several words from Lewis Carroll's 'Jabberwocky' are blends that depend in part upon phonesthesia to be understood, including *slithy* (*lithe* x *slimy*), with the pejorative *sl*- phonestheme (*sludge*, *slop*), and *mimsy* (*miserable* x *flimsy*), evocative of

whimsy and *clumsy* as well as *flimsy* (Firth 1957: 194). Firth (1930: 186) suggests that many words formally identified as blends of two lexemes are really blends of phonesthetic groups. He takes issue with Jespersen's accounting of *twirl* as a blend of *twist* and *whirl*, suggesting that 'we cannot limit the habit background of *twirl* to those two words. This background probably includes the *tw*- and *-irl/-url* phonaesthemes'. It is likely that the whole *tw*- phonesthetic group is represented in this blend, including *twist*, *twitch*, *twinge* and others – selecting only one as the definitive initial element of this blend is 'not ... a satisfactory basis' (ibid.). Algeo (1977: 60) also noted that blending may take place between 'classes of words', citing *glop*, 'a liquid or viscous substance or mixture; spec. inferior or unappetising food' (*OED*), which 'might be explained simply as a blend of *glob* and *slop*', but is more likely a blend of the *gl*-, found in *gloom*, *glug*, and *glum*, with the *-op* in *slop*, *drop* and *flop*. As discussed in section 2.2, the *OED* etymologizes *clash*, *crash* and *bash* in much the same way.

4 *Blood* and *bone*

Supporters of the Germanic Substrate Hypothesis often mention *blood* and *bone* as words with possible non-IE origins (Hawkins 1987; Boutkan & Siebenga 2005). Neither has clear non-Germanic cognates or widely accepted origins, and their proposed etymologies are semantically problematic. I suggest that the role of phonesthesia in lexical development has been overlooked, and that applying its principles to these words may be fruitful. Based on the attraction of the *b*- 'body-related' phonestheme, I suggest that *blood* and *bone* were formed as phonesthetic blends, along the same lines as Old English *docga* 'dog', as discussed in section 2.1.3.

4.1 *Blood*

Figure 1 lists some of *blood*'s cognates and provides an overview of its proposed etymologies. It is reproduced as figure 5.

Gothic *bloþ*, Old English *blōd*, Old Norse *blōð*, Dutch *bloed*, German *Blut*
Proto-Germanic **blōda*- 'blood'

Derived from Proto-Germanic **blēan*- 'to blow', relating to the notion of gushing (*OED*; Skeat 1887; Kroonen 2013), or **blōan*- 'to flower', relating to blushing (Kroonen 2013), or **blēda*- 'breath', with an association with 'life' (Kroonen 2013). Possibly related to Proto-Indo-European (PIE) **b^hleH-* 'swell, blow up, bubble' (Boutkan & Siebenga 2005). Watkins (2011) suggests PIE **bhel-* 'to thrive, bloom', suffixed form PIE **bhlō-to-* 'possibly in the meaning' 'swell, gush, spurt'. Orel (2003) provides the same analysis morphologically, glossing **bhel-* as 'to swell'. Lloyd *et al.* (1998: 211–12) largely agree, but gloss **b^hlō-* as 'swell, gush, swell'. Boutkan & Siebenga (2005) suggest that this word may not descend from PIE at all.

Figure 5. *Blood*'s etymology

None of these etymologies seems definitive. The only two things that all of the sources I've consulted agree on is that the word existed in Proto-Germanic as an *a*-stem noun, and that it is a derived form of some kind. All require some imagination, most lack an explicit explanation of how the word arrived at its Proto-Germanic form, and none is reported with particular confidence: Kroonen (2013) lists three suggestions without any mention of which one is most plausible; Watkins carefully qualifies his etymology; Boutkan doubts the PIE in his own entry; and all provide different definitions of the PIE root.

A thoughtful analysis is provided by Lloyd *et al.* (1998: 211–12). They note that the abundance of words for 'blood' in the Indo-European languages is 'usually explained by the replacement of taboo words with euphemisms'. Noting the abundant discord over *blood*'s etymology, they argue that the most likely analysis is that the Germanic etymon of *blood* (**blōda-* or **blōþa-*) is one such euphemism, probably originating as the 'perfect participle *-to-* added to the Indo-European **bhlō-* "pour, well, gush, swell"'. **blōda-* would thus mean something like 'that which gushed'. This analysis is morphologically reasonable and semantically quite transparent. It also largely accords with those provided by Orel (2003: 50) and Watkins (2011: 10).

Setting aside slight differences in their reconstructed forms, all three derive the Germanic word for 'blood' from the PIE **b^hlō*, an o-grade ablaut of **b^hel(ə)*. The significant point of departure involves the meaning of the Indo-European stem **b^hlō-*.

Lloyd *et al.* gloss **b^hlō* as 'quellen', translated to English as 'pour, well, gush, swell' (Messinger, Türk & Willmann 1993: 463). Fick *et al.* (1909: 146) earlier provided the same definition, but noted their uncertainty by appending a question mark to their gloss: 'quellen?'

Watkins, on the other hand, provides no explicit definition for **bhlō-*, leaving it to inherit its definition from its basic, e-grade form, **bhel-* 'to thrive, bloom'. Interestingly, none of the words that Watkins derives from **bhlō-* (e.g. *blow*, *bloom*, *blossom*, and, via Latin, *flower*) points to any connection with gushing, welling or pouring apart from the Germanic blood-related words. Watkins suggests that the 'suffixed form **bhlō-to-*', the etymon of these (and only these) Germanic blood-related words, may have the meaning 'swell, gush, spurt' – meaning that he holds, contrary to Lloyd *et al.*, that **bhlō-* alone does not have this meaning.

Orel (2003: 50) likewise does not identify the PIE as having anything to do with pouring or gushing. Rather, he notes only that it is 'further connected with [Germanic] **blōanan*', i.e. 'to blow, to bloom, to blossom'.

It should be noted that, although they doubt the word's Indo-European origins, Boutkan & Siebenga (2005) do refer to a PIE **b^hleH-* 'swell, blow up, bubble'. 'Bubble' is similar to 'gush, soak' in that both denote the presence of a liquid. But since the only words derived from PIE **bhel-* that explicitly have to do with liquids are those found in Germanic, and since none of them involves any liquid other than blood, it seems to me that Watkins' more conservative definition of PIE **bhel-* 'thrive, bloom' is more warranted than Lloyd *et al.*'s.

Thus, while Lloyd *et al.*'s analysis is attractive and well reasoned, there is still an unfortunate degree of uncertainty in the meaning of the PIE stem. If indeed it indicated

gushing liquids, then an analysis of **blōda* as ‘that which gushed’ is plausible. But it is problematic that the only indication that **bhlō* had this meaning is a single word controversially derived from it.

With apologies, I add to the discord. I propose that the etymon of *blood*, reconstructed by Kroonen (2013) as PGmc. **blōda-*, was formed by blending the *b-* ‘body-related’ phonestheme with the Proto-Germanic etymon of *flood*. Figure 6 lists some of *flood*’s cognates and provides a brief sketch of its etymology.

Old English *flōd*, Gothic *flodus*, Old Norse *flóð*, Dutch *vloed* ‘flood, high tide’, German *Flut* ‘river, tide’
Proto-Germanic **flōdu-* (Kroonen 2013; but see below)

A nominalized form of Proto-Germanic **flōan-* ‘to flow’, thus ‘something that flows’. Kroonen (2013) notes that ‘the *u-* suffix has been replaced by *a-* and *i-* stems in many languages’, and bases his reconstruction on the Gothic form. Orel (2003) gives **flōdan*, an *a-* stem, also derived from the Proto-Germanic verb meaning ‘to flow’. Fick *et al.* (1909) suggest an *a-* stem and *u-* stem, *flōda* and *flōdu*, as coexisting in Proto-Germanic.

Figure 6. *Flood*’s etymology

Note that *flood*’s meaning has changed substantially through the centuries. As stated above, PGmc. **flōdu-* is simply a nominalization of the verb meaning ‘to flow’; it could refer to a flood in the modern sense of the word, but also to any body of flowing water. This was still the case for Old English *flōd*. According to the *Dictionary of Old English: A to I Online* (Cameron *et al.* 2018, hereafter *DOE*), *flōd* could variously mean ‘flowing (in) of the tide’, ‘body of (flowing) waters’, ‘river, stream’, ‘sea, ocean’, ‘water (as opposed to other elements)’, ‘deluge, inundation’ and the ‘Deluge recorded in the book of Genesis’, along with figurative meanings, like ‘copious flow/stream (of blood/tears)’ and ‘a stream of words’. That Old English *blōd* is recorded as also meaning ‘vein’ is intriguing, since *flōd* could refer to water as well as the channel that carried it (Hall 1960: 52).

In combining *b-* ‘part of the body’ with **flōdu-* ‘something that flows’, one constructs a word that denotes the substance that flows through the body. I suggest that the semantics of this etymology are clearer than many of those previously cited, and largely identical with the semantics in Lloyd *et al.*’s analysis: blood is the (most salient) fluid that flows through the body. This concept surfaces even today: when we speak of the *bloodstream*, we relate blood with flowing water, and when we say *bloodflow*, we connect *blood* with a cognate of *flood*.

4.1.1 The stem problem

For the etymon of *blood*, Kroonen (2013) reconstructs PGmc. **blōda-* as an *a-* stem, yet reconstructs the etymon of *flood* as PGmc. **flōdu-*, a *u-* stem. This presents a problem for my analysis: if **blōda-* is a blend of an initial *b-* with **flōdu-*, why should the stem of the noun have changed in the process?

The explanation may be found in the way that Kroonen reconstructs **flōdu-*. He notes that Gothic *flodus* provides the basis of his reconstruction, writing that many Germanic languages have ‘replaced’ the apparently original *u*-suffix – but offers no explanation for this development, nor of why one should favor the Gothic over the Old English or Old Norse. Stem discrepancies themselves are not strange. For example, Old English *gāt* ‘goat’ and Gothic *gaitis* are inconsistent: the Gothic word suggests a PGmc. *i*-stem, reconstructed as **gaitiz*, but this would have produced Old English *gāet* and English [git]; the stem-type of the Proto-Germanic form is necessarily ambiguous (Peeters 1977: 167). In the case of *flood*, however, the cross-linguistic discrepancies are numerous. It seems likely that the variation existed in Proto-Germanic. If the Proto-Germanic form was grammatically unstable, this would explain the great variation in its child forms.

Kroonen is apparently alone in his reconstruction; I have found no independent sources that cite *flood*’s Proto-Germanic origin exclusively as a *u*-stem. Orel (2003) reconstructs **flōðan*, and Fick *et al.* (1909) reconstruct **flōda* and **flōdu* as coexisting in PIE. As shown in table 3, the stem for the etyma of *flood* and *blood* is identical in these two sources.

Table 3. *Flood and blood reconstructions*

	flood	blood
Kroonen	<i>*flōdu-</i>	<i>*blōda-</i>
Orel	<i>*flōðan</i>	<i>*blōðan</i>
Fick	<i>*flōda & *flōdu</i>	<i>*blōda</i>

If the attested Germanic cognates of *flood* and *blood* in table 4 are compared, it becomes immediately apparent that, apart from the initial *b-*, and aside from Gothic, they are identical.

Table 4. *Flood and blood by language*

	flood	blood
Old English	<i>flōd</i>	<i>blōd</i>
Gothic	<i>flodus</i>	<i>bloþ</i>
Old Norse	<i>flóð</i>	<i>blóð</i>
Dutch	<i>vloed</i>	<i>bloed</i>
German	<i>Flut</i>	<i>Blut</i>

The Gothic divergence is accounted for easily enough taking as a starting point Fick *et al.*’s (1909) analysis that **flōdu-*, a *u*-stem version of *flood*’s etymon, coexisted in Proto-Germanic with **flōda-*, an *a*-stem version of this same word. In Proto-Germanic,

it was **flōda-* that was blended with *b-* to yield **blōda-*, which became Gothic *bloþ*. Meanwhile, it was **flōdu-* that became Gothic *floodus*. **flōdu-* and **flōda-*, which coexisted in Proto-Germanic, competed for survival in the Germanic languages, leading to a noteworthy degree of variation.

4.1.2 Blood as a phonesthetic blend

In light of the above, I suggest the blending process in [figure 7](#) as responsible for *blood* and its cognates:

$$\begin{array}{rcccl}
 b- & & *flōda- & = & *blōda- \\
 \text{'body-related'} & \times & \text{'body of} & & \text{'blood'} \\
 & & \text{flowing water'} & &
 \end{array}$$

Figure 7. *Blood*

In this sort of blend, there is a motivation to preserve as much of the second element as phonologically possible. As the initial *bl-* cluster is permissible in Proto-Germanic, the *b-* phonestheme only replaces the initial *f-* of **flōda-*, and not the entire onset.

4.2 Bone

[Figure 2](#), reproduced as [figure 8](#), lists some of *bone*'s cognates and summarizes its proposed etymologies.

Old English *bān*, Old Norse *bein* 'bone, leg', Dutch *been* 'bone, leg', German *Bein* 'leg, bone (archaic)', Proto-Germanic **bainā-* 'bone, leg' (Kroonen 2013; but see below)

If connected to Old Norse *beinn* 'straight', it may be derived from Proto-Indo-European **bheh₂-* 'to shine', assuming that Proto-Germanic **bainā-* originally meant 'beam, ray, post' (Kroonen 2013). Alternatively, original meaning in Proto-Germanic may have been 'long bone of the leg' (*OED*) or just 'bone' (Urban 2015; Orel 2003; Lloyd & Springer 1988). May also come from Proto-Indo-European **bhei(ə)-* 'to strike' (Orel 2003; Lloyd & Springer 1988). Cate-Silfwerbrand (1958) argued for a Celtic loanword derived from the Proto-Celtic **bend*, **bnd-no-* 'projecting tip, horn' (Lloyd & Springer 1988: 516); Hawkins (1987: 75) claims non Indo-European etymology.

Figure 8. *Bone*'s etymology

As is the case with *blood*, there is little consensus regarding the development of *bone*. It has been characterized as 'quite isolated in Indo-European and without etymology' (Lloyd & Springer 1988: 515). A connection to Icelandic and Old Norse *beinn* 'straight' is usually mentioned, but often doubtfully (Lloyd & Springer 1988: 515). This is anecdotally supported by the fact that some south German dialects use *Bein* to

refer to long, straight bones, as opposed to rounded joints. However, this etymology is significantly problematized by the fact that the adjective *beinn* is ‘exclusive to northern Germanic’, while correspondences to *bone* are found ‘in all Germanic languages except for Gothic’. Lloyd & Springer (1988: 515–16) find this reconstruction to be ‘etymologically opaque’. The *OED* similarly notes that the connection to *beinn* ‘cannot be either substantiated or disproved, as that word is itself of uncertain origin and without parallels in West Germanic’. Even if the words shared a Proto-Germanic root, there is no compelling reason to suggest that the word for ‘bone’ is derived from the word for ‘straight’ in Proto-Germanic.

The connection to ‘to strike’ is unclear as well: words derived from PIE **bhei-* ‘to strike’ include Old Irish *benim* ‘strike, cut’, Russian *bito* ‘stick, hammer’, Armenian *bir* ‘big stick, club’ and English *bill* (of a bird) (Lloyd & Springer 1988: 516; Watkins 2011: 9). There seems to be a suggestion that bones are shaped like clubs, or got their name from having been used as a beating implement. This is possible, but not widely supported, and seems semantically uncertain.

Lloyd & Springer (1988: 515–16) briefly discuss several other, less popular etymological proposals as well, all of which they characterize as unconvincing: connections to Latin *femur* ‘thigh’, Norwegian *buna* ‘bone tube’ and Middle Low German *bunk* are rejected because these vowel forms would not have arisen from the diphthong in Proto-Germanic **baina-*, and Cate-Silfwerbrand (1958) calls upon an otherwise undocumented Celtic loanword with *i*-epenthesis, relating to Proto-Celtic (possibly PIE) **bend-*, **bnd-no-* ‘projecting tip, horn’, rejected as unclear.

In summary, there have been many proposed etymologies for **baina-*. None is generally accepted, and all are recognized as troublesome, on either formal or semantic grounds.

4.2.1 Bone’s polysemy

Because of its polysemy in the Germanic languages, Kroonen (2013) reconstructs the meaning of **baina-* as ‘bone, leg’. This is at odds with general consensus: Lloyd & Springer (1988: 515–16), Seebold (2001), Orel (2003) and Urban (2015: 385) reconstruct the original meaning as ‘bone’, while the *OED* notes that it may have ‘denoted a long bone of the leg’. There is considerable evidence that Proto-Germanic **baina-* meant ‘bone’, and that ‘leg’ was a later development. Urban (2015) bases this on many factors, including extensive internal evidence in the development of German, the surviving textual evidence, and the meaning of *bone*’s cognates in most compounds and derived forms that include it.

Old Norse, Old English and Old High German are the earliest languages with a written cognate of *bone*. No Gothic words for ‘bone’ or ‘leg’ have survived; the passages in the Bible that would have contained word for ‘bone’ are missing from Ulfila’s translation (Cleasby & Vigfusson 1874: 55). In Old Norse, *bein* primarily meant ‘bone’, but a meaning of ‘leg’, specifically from the knee to the foot, is attested in later sources (Cleasby & Vigfusson 1874: 55). The usual word for ‘leg’ in Old Norse was *leggr* (Arthur 2002: 85). Old English *bān* chiefly meant ‘bone’, but it seems to have meant

‘leg’ in certain compounds, discussed below. However, the *OED* notes that ‘the sense ‘leg’ is not unambiguously attested for the simplex in Old English’, and in any case *sceanca* – today’s *shank* – was the usual word for ‘leg’.

In none of these languages did the cognates of *bone* refer exclusively to the bones of the leg, and numerous compounds containing *bone*-cognates point to its general meaning. Consider the Old Norse *viðbeina* ‘collar-bone’ and *höfuðbein* ‘head-bones’, Old High German *brustbein* ‘breast bone’ and Old English *cinbān* ‘jawbone, jaw, chin’.

The hints at polysemy in Old English are found in some compounds, including *bānece* ‘pain in the thigh (-bone), sciatica’ and *bānriht* ‘leg armour, greave(s), literally “bone-covering” or “leg-covering”’, which was used to gloss the Latin *tibialis* (*DOE*). Like the situation in Old Norse, the Old High German *bein* originally meant ‘bone’, but later came to mean ‘leg’ as well (Urban 2015: 374). Urban explains the semantic broadening of **baina-* as an example of metonymy, ‘based on spatial contiguity’ (2015: 375). The bones of the leg may have been the most salient, being the largest in the body.

4.2.2 Bone as a phonesthetic blend

Accepting the original Proto-Germanic meaning as ‘bone’, I suggest that **baina-* is a blend of the *b-* ‘body-related’ phonestheme with Proto-Germanic **staina-* ‘stone’. Figure 9 provides some of *stone*’s cognates and an overview of its etymology.

Old English *stān*, Gothic *stains*, Old Norse *steinn*, Dutch *steen*, German *Stein*
 Proto-Germanic **staina-* ‘stone’

From Proto-Indo-European **stāi-* ‘stone’ (Watkins 2011). Kroonen (2013) lists Old Church Slavonic *stěna*, Russian *stená* ‘wall’, and Greek *stia*, *stíon* ‘small stone, pebble’ as cognates. Orel (2003) seconds the Slavic connection, providing Proto-Slavic **stěna* ‘(stone) wall’.

Figure 9. *Stone*’s etymology

Table 5 compares *bone* and *stone* and their cognates across several Germanic languages. Gothic is absent because no cognate of *bone* is attested in that language.

Table 5. *Bone* and *stone* by language

	bone	stone
Old English	<i>bān</i>	<i>stān</i>
Old Norse	<i>bein</i>	<i>steinn</i>
Dutch	<i>been</i>	<i>steen</i>
German	<i>Bein</i>	<i>Stein</i>

The Germanic peoples were intimately familiar with both bones and stones, and runic inscriptions are found carved in both materials. Stones are the hard mineral objects abundant in the natural world, and bones are the hard, seemingly mineral objects in the body. Thinking of a ‘bone’ as a ‘body-stone’ is semantically transparent – the only other real contender for this meaning would have been ‘teeth’.

I suggest the blending process in figure 10 as responsible for *bone* and its cognates. As with *blood*, as much of the second element was preserved as phonologically possible. Because an initial *bt-* cluster is illegal in Proto-Germanic, the *b-* phonestheme replaces the whole onset of the syllable.

$$\begin{array}{ccccccc}
 b- & & x & & *staina- & & = & & *baina- \\
 \text{‘body-related’} & & & & \text{‘stone’} & & & & \text{‘bone’}
 \end{array}$$

Figure 10. *Bone*

5 Conclusion

Because of a paucity of data from outside Germanic, it is impossible to know the origins of words like *blood* and *bone* with certainty. As shown above, there is very little concord on their origins. The Germanic Substrate Hypothesis has accounted for many difficult words by calling upon a hypothetical substrate language. While this method may be valid, it is also fruitful to appeal to productive processes of word development. By applying the principles of phonesthetics to the problem of *blood* and *bone*, new avenues of etymology may have been opened. I suspect that this sort of analysis may be useful in tackling other Germanic words with problematic etymologies; building on the work of Gąsiorowski (2006), it can explain the origin of *dog*, and Bolinger and Firth have used it to account for many other lower-frequency words.

In addition to providing new etymologies for *blood* and *bone*, this proposal also incorporates phonesthemes into etymological work in a new way. As shown in prior sections, English words like *twiddle*, *sneeze*, *crash*, *clash*, *bash*, *frog*, *bug*, *twirl* and *glop* have been described by earlier researchers as owing their forms or meanings to phonesthesia. For many of these words, the effect of phonesthesia has been relatively small – the modification of a consonant in a pre-existing word, as with *sneeze* and *bug*, or a shift of meaning, as with *twiddle*.

In this proposal, meanwhile, phonesthemes are recognized not merely as strings which might prompt the modification of an existing word, but as word-formation elements in their own right. This itself is not entirely new: some words, like *clash* and *twirl*, have been etymologized as blends or compounds of phonesthemes before, and, as shown in section 2.1.3, *dog* may also be analyzed as the blending of a word (*dox* ‘yellow-brown’) with a phonestheme (*-gga* ‘hypocoristic animal name’). As Kwon & Round (2015: 24) highlight, most phonesthemes diverge from traditional morphemes in that they often appear in ‘lexical stems which are composed of a recurring, sound-meaning pairing [i.e. the phonestheme itself] plus a non-recurrent residue’. It

follows that phonesthemes should have played a role in the development of these stems, and therefore in the creation of wholly new words. What is new in this proposal, however, is the suggestion that phonesthemes may have been essential to the composition of high-frequency words.

The present sense generally seems to be that phonesthesia, where recognized, is a marginal phenomenon, and thus must appear only in novel, low-frequency words. However, this attitude is not universal. Joseph (1997: 15) wrote of the importance of ‘marginal’ linguistic phenomena to a systematic understanding of grammar, noting that, for the speaker, ‘all linguistic knowledge starts out as marginal’, and that ‘generalizations’ emerge as we learn more and fit these ‘marginal’ scraps of linguistic information into a system. He goes on to note that ‘the smaller, more local, generalizations are what speakers exploit dynamically, as the case of phonesthematic attraction shows’.

Joseph presaged a trend of interest in ‘marginal’ linguistic phenomena, and his paper has been cited in over a dozen articles on topics such as ideophones, sound symbolism and phonesthesia.⁷ In her monograph *Rhyme over Reason*, Réka Benczes (2019: 18) takes note of the ‘growing body of research into atypical patterns in English word-formation’, and contributes by exploring words formed through phonological motivation. This includes phonesthematic (‘sound-symbolic’) words such as ‘*glimmer* and *glisten*’ and blends, along with words formed through wordplay and alliterative compounding. Benczes maintains that words formed through phonological motivation are ‘part and parcel of everyday language use’. Even if phonesthesia is a less frequently employed method of word formation, there is no reason to suppose that the words it is used to coin must always remain low-frequency, novelty words. For example, the *OED* places *clash*, *crash* and *frog* as ‘occur[ing] between 1 and 10 times per million words in typical modern English usage’, on par with words like *surveillance*, *tumult* and *paraphrase*.

It is well established that innovative words have historically replaced more staid synonyms. Perhaps the most popular example is that of the Germanic words for ‘bear’. Owing to a taboo on saying the animal’s name, there is no Germanic animal word derived from the PIE **rtko-* ‘bear’. Instead, Proto-Germanic **beran-* ‘bear’ was derived from the PIE **bher-* ‘bright, brown’, allusively referring to the animal as something like ‘the brown one’ (Watkins 2011: 74). Another case is that described in section 2.1.3, where the hypocoristic OE *frogga* came to replace *frosco/frox*. Like these words, *blood* and *bone* must each have coexisted with earlier words for the concepts that they refer to. PIE roots with these meanings have been reconstructed and, with some variation in specifics, are widely accepted. No words derived from PIE **ost-* ‘bone’ are found in Proto-Germanic (Orel 2003; Watkins 2011: 63; Kroonen 2013), but PIE **kreuə-* ‘blood’ gave rise to Proto-Germanic **hrawa-* ‘raw’ and **hraiwa-* ‘dead body, death’ (Kroonen 2013). Words derived from these roots survived in the modern

⁷ See, for example, Smith (2016), Alderete & Kochetov (2017) and Dingemanse (2018).

Germanic languages, including, for example, English *raw* from the former and Dutch *reeuw* ‘foam or sweat of a dying person’ from the latter (Kroonen 2013). But over time, these newer phonesthetic formations – **blōda* and **baina* – came to be favored, and replaced the pre-existing words for ‘blood’ and ‘bone’.

Author's address:

Linguistics MA/PhD Program
The Graduate Center, CUNY
Room 7407 (mailbox 78)
365 Fifth Ave
New York, NY 10016
USA
joseph.pentangelo@gmail.com

References

- Abelin, Åsa. 1999. *Studies in sound symbolism* (Gothenburg Monographs in Linguistics 17). PhD dissertation, University of Gothenburg.
- Abelin, Åsa. 2015. Phonaesthemes and sound symbolism in Swedish brand names. *Ampersand* 2, 19–29.
- Alderete, John & Alexei Kochetov. 2017. Integrating sound symbolism with core grammar: The case of expressive palatalization. *Language* 93(4), 731–66.
- Algeo, John. 1977. Blends, a structural and systemic view. *American Speech* 52(1/2), 47–64.
- Arthur, Ross G. 2002. *English–Old Norse dictionary*. Cambridge, ON: In Parentheses Publications.
- Baldi, Philip & B. Richard Page. 2006. Review of *Europa Vasconica–Europa Semitica* by Theo Vennemann. *Lingua* 116, 2183–220.
- Bat-El, Outi. 2013. Blends. In Geoffrey Khan (ed.), *Encyclopedia of Hebrew language and linguistics*, vol. 1, 371–3. Leiden: Brill.
- Bauer, Laurie. 2006. Compounds and minor word-formation types. In Bas Aarts & April McMahon (eds.), *The handbook of English linguistics*, 483–506. Hoboken, NJ: Blackwell.
- Benczes, Réka. 2019. *Rhyme over reason: Phonological motivation in English*. Cambridge: Cambridge University Press.
- Bergen, Benjamin K. 2004. The psychological reality of phonaesthemes. *Language* 80(2), 290–311.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.
- Blust, Robert. 2003. The phonestheme ŋ- in Austronesian languages. *Oceanic Linguistics* 42(1), 187–212.
- Blust, Robert. 2011. The problem of doubleting in Austronesian languages. *Oceanic Linguistics* 50(2), 399–457.
- Bolinger, Dwight. 1940. Word affinities. *American Speech* 15(1), 62–73.
- Bolinger, Dwight. 1953. The life and death of words. *The American Scholar* 22(3), 323–35.
- Bolinger, Dwight. 1965. *Forms of English: Accent, morpheme, order*. Cambridge, MA: Harvard University Press.
- Bolinger, Dwight. 1975. *Aspects of language*, 2nd edn. New York: Harcourt Brace Jovanovich.

- Boussidan, Armelle, Eyal Sagi & Sabine Ploux. 2009. Phonaesthetic and etymological effects on the distribution of senses in statistical models of semantics. *Proceedings of the 34th Cognitive Science Annual Meeting*, 36–40.
- Boutkan, D. F. H. 1998. On the form of North European substratum words in Germanic. *Historische Sprachforschung / Historical Linguistics*, 111(1), 102–33.
- Boutkan, Dirk & Sjoerd Michiel Siebenga. 2005. *Old Frisian etymological dictionary*. <http://dictionaries.brillonline.com/frisian> (accessed 18 August 2019).
- Burridge, Kate & Tonya N. Stebbins. 2015. *For the love of language*. Cambridge: Cambridge University Press.
- Cameron, Angus, Ashley Crandell Amos & Antonette diPaolo Healey (eds.). 2018. *Dictionary of Old English: A to I online*. <http://tapor.library.utoronto.ca/doe/index.html> (accessed 18 August 2019).
- Carling, Gerd & Niklas Johansson. 2014. Motivated language change: Processes involved in the growth and conventionalization of onomatopoeia and sound symbolism. *Acta Linguistica Hafniensia* 46(2), 199–217.
- Cate-Silfwerbrand, Ragnhild Birgitta ten. 1958. *Vlees, bloed en been: Synoniemvergelijkend onderzoek van drie Germaanse woordformaties*. Assen: Van Gorcum.
- Cleasby, Richard & Gudbrand Vigfusson. 1874. *An Icelandic–English dictionary*. Oxford: Oxford University Press.
- Dingemans, Mark. 2018. Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A journal of general linguistics* 3(1), 4.1–30. <https://doi.org/10.5334/gjgl.444> DOE. See Cameron *et al.* (2018) above.
- Elsen, Hilke. 2017. The two meanings of sound symbolism. *Open Linguistics* 3(1), 491–99. <https://doi.org/10.1515/opli-2017-0024>
- Feist, Sigmund. 1914. *Indogermanen und Germanen: Ein Beitrag zur Europäischen Urgeschichtsforschung*. Halle: Niemeyer.
- Feist, Sigmund. 1932. The origin of the Germanic languages and the Indo-Europeanising of North Europe. *Language* 8(4), 245–54.
- Fick, August, Hjalmar Falk & Alf Torp. 1909. *Wörterbuch der Indogermanischen Sprachen: Dritter Teil: Wortschatz der Germanischen Spracheinheit*. Göttingen: Vandenhoeck und Ruprecht.
- Findell, Martin. 2012. *Phonological evidence from the continental Runic inscriptions*. Berlin: De Gruyter Mouton.
- Firth, J. R. 1930. *Speech*. London: Benn's Sixpenny Library. Reprinted in Peter Strevens (ed.), 1964, *The tongues of men and speech*, Oxford: Oxford University Press.
- Firth, J. R. 1957. *Papers in linguistics, 1934–1951*. Oxford: Oxford University Press.
- Fortson, Benjamin W. 2010. *Indo-European language and culture*. Chichester: Wiley-Blackwell.
- Francis, W. Nelson & Henry Kucera. 1982. *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Gąsiorowski, Piotr. 2006. The etymology of Old English **docga*. *Indogermanische Forschungen* 111, 275–84.
- Gollancz, Israel (ed.). 1895. *The Exeter Book, an anthology of Anglo-Saxon poetry*. London: Early English Text Society.
- Gries, Stefan. 2004. Shouldn't it be breakfunch? A quantitative analysis of blend structure in English. *Linguistics* 42(3), 639–67.
- Hall, J. R. Clark. 1960. *A concise Anglo-Saxon dictionary*. Toronto: University of Toronto Press.
- Hawkins, John A. 1987. Germanic languages. In Bernard Comrie (ed.), *The world's major languages*, 69–76. Oxford: Oxford University Press.
- Hogg, Richard M. 1982. Two geminate consonants in Old English? In John A. Anderson (ed.), *Language form and linguistic variation: Papers dedicated to Angus McIntosh*, 187–202. Amsterdam: John Benjamins.

- Hough, Carole & Christian Kay (eds). 2017. *Thesaurus of Old English (TOE) online*. <https://oldenglishteaching.arts.gla.ac.uk/> (accessed 29 October 2019).
- Jones, Sir William. 1798. Third anniversary discourse: On the Hindus. *Asiatick Researches* 1, 415–31.
- Joseph, Brian D. 1997. On the linguistics of marginality: The centrality of the periphery. www.asc.ohio-state.edu/joseph.1/publications/1997_ont.pdf (accessed 18 August 2019).
- Kroonen, Guus. 2011. *The Proto-Germanic n-stems: A study in diachronic morphology*. Amsterdam: Rodopi.
- Kroonen, Guus. 2012. Non-Indo-European root nouns in Germanic: Evidence in support of the agricultural substrate hypothesis. In Riho Grünthal & Petri Kallio (eds.), *A linguistic map of prehistoric Northern Europe*, 239–60. Helsinki: Suomalais-Ugrilainen Seuran Toimituksia.
- Kroonen, Guus. 2013. *Etymological dictionary of Proto-Germanic*. <http://dictionaries.brillonline.com/proto-germanic> (accessed 16 August 2019).
- Kwon, Nahyun & Erich R. Round. 2015. Phonaesthemes in morphological theory. *Morphology* 25, 1–27.
- Lawler, John. 2006. The data fetishist's guide to rime coherence. www.umich.edu/~jlawler/rimecoherence.pdf (accessed 19 August 2019).
- Lehmann, Winfried P. 1971. *The development of Germanic verse form*. New York: Gordian Press.
- Liberman, Anatoly. 2006. Review of Dirk Boutkan and Sjoerd Michiel Siebinga, *Old Frisian Etymological Dictionary. Trefwoord*. https://ivdnt.org/images/stories/onderzoek_en_onderwijs/publicaties/trefwoord/oeftd-review.pdf (accessed 18 August 2019).
- Liberman, Anatoly. 2008. *An analytic dictionary of English etymology: An introduction*. Minneapolis: University of Minnesota Press.
- Lloyd, Albert L. & Otto Springer. 1988. *Etymologisches Wörterbuch des Althochdeutschen*, vol. 1. Göttingen and Zurich: Vandenhoeck & Ruprecht.
- Lloyd, Albert L., Rosemarie Lühr & Otto Springer. 1998. *Etymologisches Wörterbuch des Althochdeutschen*, vol. 2. Göttingen and Zurich: Vandenhoeck & Ruprecht.
- Magnus, Margaret. 1998. *Phonosemantic dictionary*. www.trismegistos.com/Dictionary.html (accessed 18 August 2019).
- Magnus, Margaret. 2013. A history of sound symbolism. In Keith Allen (ed.), *The Oxford handbook of the history of linguistics*, 191–208. Oxford: Oxford University Press.
- Messinger, Heinz, Gisela Türck & Helmut Willmann. 1993. *Langenscheidt's standard German dictionary*. New York, Berlin, Munich, Vienna and Zurich: Langenscheidt.
- OED = Oxford English dictionary online* (2018). www.oed.com (accessed 18 August 2019).
- Orel, Vladimir. 2003. *A handbook of Germanic etymology*. Leiden: Brill.
- Owen, Francis. 1928. Alliteration in the runic inscriptions. *Modern Philology* 25(4), 397–408.
- Peeters, Christian. 1977. Notes on Germanic etymologies. *Zeitschrift für vergleichende Sprachforschung* 91(1), 166–9.
- Pereltsvaig, Asya. 2010. Blending in. *Languages of the world*. www.languagesoftheworld.info/morphology/blending-in.html (accessed 18 August 2019).
- Pereltsvaig, Asya & Martin Lewis. 2015. *The Indo-European controversy*. Cambridge: Cambridge University Press.
- Philps, Dennis. 2011. Reconsidering phonaesthemes: Submorphemic invariance in English 'sn-words'. *Lingua* 121, 1121–37.
- Prokosch, Eduard. 1939. *A comparative Germanic grammar*. Baltimore, MD: Linguistic Society of America.
- Roberge, Paul. 2010. Contact and the history of Germanic languages. In Raymond Hickey (ed.), *The handbook of language contact*, 406–31. Hoboken, NJ: Wiley-Blackwell.
- Sadowski, Piotr. 2001. The sound-symbolic quality of word-initial 'gr-' cluster' in Middle English alliterative verse. *Neuphilologische Mitteilungen* 102(1), 37–47.

- Schuhmann, Roland. 2012. Where is the substrate in the Germanic lexicon? www.academia.edu/1928440/Where_is_the_Substrate_in_the_Germanic_Lexicon (accessed 18 August 2019).
- Seebold, Elmar. 2001. *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben*. The Hague: Mouton.
- Skeat, Rev. Walter, W. 1887. *Principles of English etymology*, vol. I: *The native element*. Oxford: Clarendon Press.
- Smith, Chris. 2014. The phonesthetics of blends: A lexicographic study of cognitive blends in the OED. *Explorations in English Language and Linguistics* 2(1), 12–45.
- Smith, Chris A. 2016. Tracking semantic change in *fl-* monomorphemes in the *Oxford English Dictionary*. *Journal of Historical Linguistics* 6(2), 165–200.
- Sörjö, Aljoša. 2015. Pre-Germanic: A tentative description of the substratum language in Proto-Germanic based on Guus Kroonen's etymological dictionary of Proto-Germanic. PhD dissertation, University of Ljubljana.
- Szymanek, Bogdan. 2005. The latest trends in English word formation. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, 429–48. Dordrecht: Springer.
- Tappenden, Graham. 2009. Teuro. ww.allthingsgerman.net/blog/podcasts/teuro/ (accessed 18 August 2019).
- Urban, Matthias. 2015. Lexical semantic change and semantic reconstruction. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 374–92. London: Routledge.
- Vennemann, Theo. 2003. *Europa Vasconica – Europa Semitica* (Trends in Linguistics, Studies and Monographs 138). Berlin and New York: De Gruyter Mouton.
- Vennemann, Theo. 2012. *Germania Semitica* (Trends in Linguistics, Studies and Monographs 259). Berlin: De Gruyter Mouton.
- Wallis, John. 1653. *Grammatica Linguae Anglicanae*. Oxford: Leon Lichfield.
- Watkins, Calvert. 2011. *The American heritage dictionary of Indo-European roots*, 3rd edn. Boston: Houghton Mifflin Harcourt.
- Williams, Graham. 2013. Glossing over the lamb: Phonaesthetic *gl-* in Middle English and aural scepticism in *Pearl*. *The Review of English Studies* 65(271), 596–618.
- Wright, David. 2012. Scrunch, growze, or chobble? Investigating regional variation in sound symbolism in the *Survey of English Dialects*. Leeds Working Papers in Linguistics and Phonetics 17.