A faster path between meaning and form? Iconicity facilitates sign recognition and production in British Sign Language

Vinson, David; Thompson, Robin L.; Skinner, Robert; Vigliocco, Gabriella

DOI:
10.1016/j.jml.2015.03.002

Journal of Memory and Language 82 (2015) 56–85

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

A faster path between meaning and form? Iconicity facilitates sign recognition and production in British Sign Language

David Vinson [a,b,*], Robin L. Thompson [a,c], Robert Skinner [a,d], Gabriella Vigliocco [a,b]

[a] Deafness, Cognition and Language Research Centre, University College London, United Kingdom
[b] Institute for Multimodal Communication, University College London, United Kingdom
[c] School of Psychology, University of Birmingham, United Kingdom
[d] The Centre for Translation & Interpreting Studies in Scotland, Heriot-Watt University, United Kingdom

ARTICLE INFO

Article history:
Received 21 January 2014
revision received 6 March 2015

Keywords:
Lexicon
Iconicity
Embodiment
Language comprehension
Language production

ABSTRACT

A standard view of language processing holds that lexical forms are arbitrary, and that non-arbitrary relationships between meaning and form such as onomatopoeias are unusual cases with little relevance to language processing in general. Here we capitalize on the greater availability of iconic lexical forms in a signed language (British Sign Language, BSL), to test how iconic relationships between meaning and form affect lexical processing. In three experiments, we found that iconicity in BSL facilitated picture–sign matching, phonological decision, and picture naming. In comprehension the effect of iconicity did not interact with other factors, but in production it was observed only for later-learned signs. These findings suggest that iconicity serves to activate conceptual features related to perception and action during lexical processing. We suggest that the same should be true for iconicity in spoken languages (e.g., onomatopoeias), and discuss the implications this has for general theories of lexical processing.

Introduction

Approaches to lexical processing, whether in comprehension or production, embed a core assumption that the relationship between word meaning and form is arbitrary. That is, words' phonology and orthography do not reflect sensorimotor characteristics of the words' referents; the existence of entirely different words across languages for the same entity being testament to arbitrariness as a general principle of linguistic form. As a result, models of lexical representation and processing assume a strong distinction between meaning and form. This separability is typically achieved via intermediate representations abstracted away from meaning which serve to bind the two domains (e.g. *lemma* in models of lexical retrieval in language production, e.g. Levelt, 1989; Levelt, Roelofs & Meyer; 1999; Dell, 1986; *lexical nodes* in models of comprehension, e.g., Norris & McQueen, 2008; Norris, McQueen & Cutler, 2000). However, the mapping between word meaning and form is not always arbitrary; and instances of meaningful, non-arbitrary (iconic) links between form and meaning such as onomatopoeia can be found (the use of speech sounds to mimic natural noises like "meow" and "quack").

While such iconic relationships appear very limited in English, they are far more prevalent in many non Indo-European languages which have extensive repertoires of sound-symbolic words referring to sensory, motor and affective experience well beyond the limited extent of the acoustic domain covered by onomatopoeias (see Perniss, Thompson & Vigliocco, 2010). Moreover, iconic relationships between meaning and form are far

* Corresponding author at: Institute for Multimodal Communication, 26 Bedford Way, London WC1H 0AP, United Kingdom.
E-mail address: d.vinson@ucl.ac.uk (D. Vinson).

http://dx.doi.org/10.1016/j.jml.2015.03.002
0749-596X/

more pervasive in signed languages, where the visual-manual modality affords many more opportunities for non-arbitrary mappings (e.g. Taub, 2001). Observations along these lines led Perniss, Thompson and Vigliocco (2010) and Perniss and Vigliocco (2014) to argue that the focus on arbitrariness at the expense of iconicity is essentially a historical accident due to exclusive focus on languages with impoverished iconic vocabularies (i.e. English and other Indo-European languages). Perniss et al. (2010) present an alternative framework, by which arbitrariness and iconicity both play foundational roles in language: arbitrariness serving communicative effectiveness by permitting maximum discrimination among the forms of entities similar in meaning (e.g. Monaghan & Christiansen, 2006; Monaghan, Christiansen & Fitneva, 2011), and iconicity linking linguistic form to human experience in the form of perception and action (e.g.; Meteyard & Vigliocco, 2008; Meteyard, Rodriguez Cuadrado, Bahrami & Vigliocco, 2012). Under this framework arbitrariness ensures fine-grained vocabulary distinctions while iconicity ensures referential links between linguistic labels and conceptual referents in development and processing. In processing it can be seen as a mechanism by which sensory-motor systems are engaged, as argued by embodied cognition views (e.g. Barsalou, 1999, 2009; Fischer & Zwaan, 2008; Gallese & Lakoff, 2005; Glenberg & Kaschak, 2003; Pulvermüller, 1999; Vigliocco, Meteyard, Andrews & Kousta, 2009).

However, no role for iconicity is assumed by standard accounts of language comprehension and production. Under such views, iconicity would not affect processing, as the abstract linkage between meaning and form strips away the details of perceptual experience even for highly iconic words (e.g., onomatopoeia) and signs (e.g., the sign HAMMER across many sign languages visually resembles the act of hammering) that would be processed as if their forms were arbitrary. Under embodied views, iconicity could be accommodated in comprehension, in terms of boosting activation of specific sensory-motor properties that are represented in the linguistic form, and in production, by boosting activation of phonological features that correspond to sensory-motor properties (Perniss & Vigliocco, 2014). However, at this stage no detailed embodied account of lexical processing has been put forward (although see; Fischer & Zwaan, 2008; Glenberg, Witt & Metcalfe, 2013).

Thus assessing whether iconicity affects processing is a test pitting the most abstract, amodal theories of processing against the most embodied ones. But more importantly, addressing the conditions under which iconicity may affect processing will provide novel constraints on how to explicitly implement the interface between lexical and conceptual information.

Here, we address these questions by investigating lexical processing in British Sign Language (BSL). Studies addressing sign languages seem especially appropriate since in these languages iconicity is far more pervasive, yet iconicity has traditionally been argued not to play a role either in language processing or language development (e.g. Orlansky & Bonvillian, 1984).

Iconic relationships between word meaning and word form

We follow Perniss et al. (2010) in defining iconicity as "regular correspondences between form and meaning... motivated by perceptuo-motor properties of real world experience" (p. 2). We thus focus here upon cases in which features of meaning derived from perceptual/motor experience are represented in a lexical form. For spoken languages, onomatopoeia is perhaps the most salient example of an iconic link between meaning and form: words referring specifically to acoustic experience like *oink*, *baa*, *clink*, *thud*, *whir*. Although iconic links in languages like English appear mainly limited to onomatopoeic words, many other languages exhibit regular mappings between meaning and form referring not only to sounds but also to sensory, motor or affective experience (e.g. Japanese: Hamano, 1998; Kakehi, Tamori & Schourup, 1996 as well numerous other languages; see Perniss et al., 2010 for a review). Certain speech sounds also appear to be motivated by specific aspects of sensory-motor experience across languages, such as the correspondence between speech sounds like (/k/, /t/) and jagged shapes on one hand, and sounds like (/b/, /l/, /m/) and rounded shapes on the other (Köhler, 1929; Ramachandran & Hubbard, 2001).

Even for spoken languages with larger iconic lexicons, the extent of non-arbitrary links between words' meanings and forms is limited, perhaps because the acoustic medium of spoken words does not map well onto meaning for many domains of knowledge. For signed languages, instead, the visual-manual modality offers the possibility of encoding far more properties of meaning in the sign's form (Taub, 2001). Across signed languages a high proportion of lexical signs are iconic, and iconicity is not only prevalent in signs referring to concrete objects and actions, but also in abstract domains like cognition, emotion, and communication. Moreover, widespread effects of iconicity in sign languages go well beyond lexical forms. For example, classifiers are used to encode motion in space, with handshapes representing objects, moving in ways that map onto the real-world motion (e.g. Emmorey, 2002; Schembri, 2003). Signed languages also make extensive iconic use of space, such as reflecting literal arrangements of physical scenes, or representing persons or objects in space in order to express more abstract relations among them iconically (e.g. Perniss, 2012).

Iconic signs are, however, conventionalised and thus can be realized in different manners across different signed languages (Klima & Bellugi, 1979); phonological parameters combine according to the phonotactic constraints of any one particular sign language, and while iconicity may be represented in some aspects of the phonological representation of any one sign, other features can be arbitrary. For example, the sign CAT[1] is iconic in both British Sign Language (BSL) and American Sign Language (ASL), with signs in both languages indicating a cat's whiskers. But, as illustrated in Fig. 1, this iconicity is realized differently in the two languages: the BSL sign uses open hands radiating

[1] Signs are represented here as English glosses in capital letters.

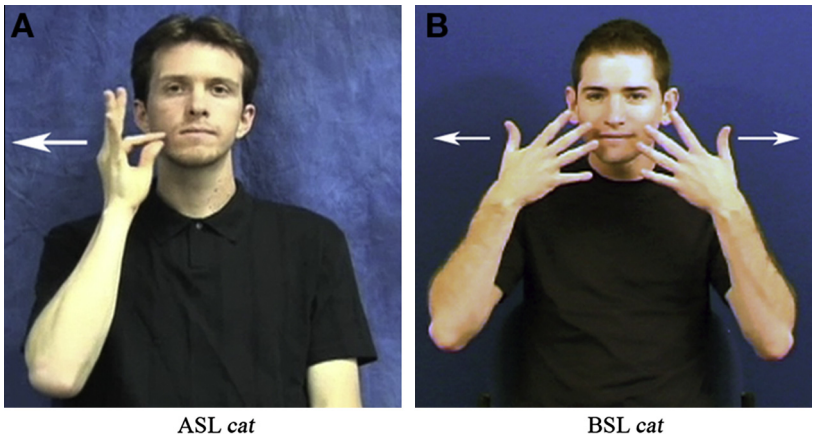

Fig. 1. Still images from the sign CAT in American Sign Language (left) and British Sign Language (right). Both signs are iconic of a cat's whiskers but differ in phonological form. The sign in ASL is also specified for two hands, but one hand is frequently dropped as shown in the figure. Figure reprinted from Perniss et al. (2010), iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, doi: http://dx.doi.org/10.3389/fpsyg.2010.00227 (copyright of the authors).

outward from the sides of the face, while the ASL sign uses a different handshape with the thumb and forefinger pressed together. This example illustrates how iconicity is combined with arbitrariness even in languages (signed languages) that use a large iconic inventory.

Implications of iconicity for models of lexical representation and processing

In order to understand how iconicity could affect lexical processing we need to make explicit assumptions about how meaning and phonology should be represented. Here we start from the generally agreed upon assumption that word meanings are componential in nature, based upon (non-linguistic) featural representations, many of which are derived from perception and action (e.g., Vigliocco, Vinson, Lewis & Garrett, 2004). Approaches of this nature have been pervasive since classical times (e.g. Aristotle's Categories, 350 BCE/1941), and modern views owe a great deal to seminal work by Rosch and colleagues in the 1970s (Rosch, 1973; Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Similarly, most models assume a phonological level, also featurally specified. In spoken languages these tend to be described in terms of articulatory mechanics (e.g. voicing, aspiration, tongue movement; Halle, 1991; Roca, 2003). Sign phonology is also described in terms of articulatory mechanics, including three major parameters (handshape, location and movement; see Sandler & Lillo-Martin, 2006 for an overview; although see Armstrong & Wilcox, 2009; Stokoe, 1991 for arguments against evaluating phonological features in sign languages without considering their iconic characteristics). In processing of both spoken and signed languages, there is substantial evidence that retrieval of phonological form can be distinguished from retrieval of meaning. For example, spontaneous errors in speaking and in signing can show either meaning or form similarity to the intended target (speech: Fromkin, 1973; Garrett, 1976, 1980, 1992; sign: Hohenberger, Happ & Leuninger, 2002; Hohenberger & Waleschkowski, 2005; Klima & Bellugi, 1979; Newkirk et al., 1980). Similarly, tip-of-the-tongue states in spoken languages reflect access to meaning but inability to retrieve a word's form (Brown, 1991; Brown & McNeill, 1966; Vigliocco, Antonini & Garrett, 1997; Vigliocco, Vinson, Martin & Garrett, 1999), and signers experience similar "tip of the fingers" states (Thompson, Emmorey & Gollan, 2005). Such evidence is consistent with views in which retrieval of wordform is separable from retrieval of semantic features.

Under these assumptions, iconic mappings between meaning and form, whether in spoken or signed language, can be characterized by a resemblance relationship between one or more non-linguistic feature of meaning, and one or more phonological feature (e.g. Kita, 1997, 2001; Taub, 2001; Vigliocco & Kita, 2006). For example, the visual appearance of the shape of an eagle's beak, and its position on the eagle's face are iconically linked to phonological features: the sign EAGLE is produced with a crooked index finger which taps the nose, and thus the iconic features are handshape (representing the shape of the beak) and location (representing the location of the beak) but not movement (as a tap is not iconic of an eagle's beak).

Standard theories of language processing (e.g., Levelt, Roelofs & Meyer; 1999; Dell, 1986; Norris et al., 2000; Norris & McQueen, 2008; Rogers et al., 2004; Vigliocco et al., 2004) have neglected iconicity and implement arbitrary meaning-form relationships via amodal intermediate representations (see Dove, 2013).

Does iconicity affect language processing?

Evidence from spoken language

A variety of studies have established that speakers of various languages are sensitive to iconicity; for example, the widespread convergence between speakers of different languages in associating certain meaningful distinctions with distinctions in word forms (assigning labels like "takete" or "kiki" for jagged shapes; labels like "maluma"

or "bouba" for rounded shapes: Aveyard, 2012; Bremner et al., 2013; Davis, 1961; Holland & Wertheimer, 1964; Köhler, 1929; Nielsen & Rendall, 2011; Ramachandran & Hubbard, 2001; Thompson & Estes, 2011; and similar studies linking properties of vowels to size of referents: Johnson, 1967; Newman, 1933; Pitcher, Mesoudi & McElligott, 2013; Sapir, 1929; Tarte and Barritt, 1971; Taylor, 1963). And similarly, illustrations that speakers can recognize aspects of meaning from unfamiliar languages (Iwasaki, Vinson & Vigliocco, 2007; Kunihira, 1971; Oda, 2000; Parault & Schwanenflugel, 2006) also indicate that indeed, some aspects of meaning are represented iconically in spoken languages and that speakers are able to recognize that those correspondences exist. Iconicity has also been shown to facilitate learning both in children (e.g. Imai, Kita, Nagumo & Okada, 2008; Kantartzis, Imai & Kita, 2011; Namy, Campbell & Tomasello, 2004; Parault & Parkinson, 2008) and adults (e.g. Kovic, Plunkett & Westermann, 2010; Monaghan, Mattock & Walker, 2012; Nygaard, Cook & Namy, 2009).

There are only a very few studies concerning the effects of iconicity in online processing of spoken languages. One recent study by Meteyard, Stoppard, Snudden, Cappa and Vigliocco (2015) showed that aphasic English speakers performed better in a reading and lexical decision task when presented with onomatopoeic words than control words matched on a number of dimensions to the onomatopoeias. In Japanese, Ohtake and Haryu (2013) found that the correspondence between vowels and size affected speeded classification of the sizes of visually presented disks. The scant evidence in this area may be linked to the limited repertoire of iconic forms in spoken Indo-European languages (those most studied in psycholinguistics).

Evidence from sign language

It has been shown that iconicity affects metalinguistic judgements of sign similarity by adults (e.g., Vigliocco, Vinson, Woolfe, Dye & Woll, 2005) but it remains controversial whether iconicity affects online processing. One line of evidence comes from picture–sign matching tasks, such as a study conducted in ASL by Thompson, Vinson and Vigliocco (2009; see also Grote & Linz, 2003). Two different kinds of pictures corresponding to the same iconic signs were used, for example, given the sign CAT (see Fig. 1) there was one picture in which the iconic property of the sign was made salient (e.g., a cat's face with its whiskers prominently displayed), and one in which the iconic property was not salient (e.g., a cat with its whiskers visible but not prominent). This task does not include an explicit judgement about the iconic property of the sign, but instead activates it implicitly by making it salient in the picture. Deaf native signers were faster in responding to a sign after seeing the salient picture (e.g., eagle with a prominent beak) than the non-salient picture (eagle with large wings), while English speakers matching the same pictures to English words showed no such effect. The authors concluded that iconicity facilitated language processing in ASL. Similar results were found for deaf children aged 8–12: iconic signs in Sign Language of the Netherlands facilitated responses in picture–sign matching, with more highly iconic signs eliciting faster and more accurate responses than signs with less iconicity (Ormel, Hermans, Knoors & Verhoeven, 2009; Ormel et al., 2010).

These processing studies provide initial evidence that iconic mappings between meaning and form can aid language processing. However, iconicity effects in picture–sign matching may come about from the use of strategies that focus attention specifically on the meaning. For example, Emmorey (2014) suggests that such iconicity effects do not reflect correspondence between lexical forms and world experience, but instead the extent to which two representations correspond with each other in line with structure-mapping theory (e.g. Gentner, 1983). Under this view, different mental representations can be aligned with each other when there is a one-to-one mapping between elements in the two domains, and the structure of relationships among elements is preserved. In the case of iconicity in signs, one-to-one mapping is reflected in features of form that correspond to features of meaning; degree of alignment between those features of form and a visually-presented stimulus would be responsible for iconicity effects. For example, the facilitation observed by Thompson et al. (2009) might be achieved because a picture expressing an iconic feature more saliently (e.g. a cat with its whiskers prominently visible) maps onto more properties of the sign form than a picture that does less so. If so, one might expect iconicity only to affect tasks where there is the possibility for structure mapping. This is particularly the case as the iconic features expressed in the ASL signs used by Thompson et al. (2009) appear to be not only iconic but also highly typical (salient to signers and non-signers alike), making it even more plausible that signers might notice the manipulation and adopt an attention-based strategy. If so, these results would not provide clear evidence for automatic activation of those semantic features iconically represented in the sign form.

Not all studies show that iconicity facilitates language processing. Thompson, Vinson and Vigliocco (2010) found an interfering effect of iconicity in a phonological decision task in which BSL signers and nonsigners were asked to indicate whether a sign had a straight or curved handshape. Although the task did not require access to meaning, the relative iconicity of the signs slowed signers' responses (differences in handshape complexity and sign production time were taken into account). This interfering effect seems at first to be at odds with the hypothesis that iconicity facilitates the mapping between form and meaning. However, as discussed in Thompson et al. (2010), this can be accounted for in terms of automatic activation of semantic features for iconic signs, features that are irrelevant for the task at hand. Iconicity also seems to have an interfering role in translation tasks. Baus, Carreiras and Emmorey (2012) asked participants to translate ASL verbs (either iconic or non iconic) in English. Proficient bilinguals were slower for iconic signs, while iconicity instead facilitated performance for non-signers who were taught those same signs during a 30-min experimental session before being tested. Baus et al. suggest that for fluent signers, iconic signs automatically involve retrieval of semantic representations, hence they would undergo deeper (conceptual) processing than noniconic signs (consistent with the results from Thompson et al., 2010). Adults learning to sign

for the first time instead may benefit from iconicity when beginning to learn a signed language, perhaps because of recruitment of their existing repertoire of iconic gestures (see Ortega & Morgan, 2010).

The above studies suggest that iconicity affects language processing by strengthening links between meaning and form. But not all studies show effects of iconicity. Bosworth and Emmorey (2010) conducted a lexical decision study using iconic ASL signs, and found that semantic priming was no greater for iconic primes than for non-iconic primes. If iconicity affects processing generally, they argued, iconic properties of primes should especially activate those semantic features that are expressed iconically, which should enhance facilitation when prime and target are both iconic. Specifically, in addition to the spreading of activation between prime and target concepts argued to be the general mechanism of semantic priming (e.g. Collins & Loftus, 1975), spreading of activation from one iconic feature to another should enhance priming when the prime and target are both iconic, regardless of whether or not they share a specific iconic property. Bosworth and Emmorey interpreted this finding as evidence that iconicity *per se* does not affect lexical processing, that is, activation of a semantic feature by an iconic sign does not lead to increased activation of other semantic features or spreading activation in a lexical-semantic network.

Furthermore, there is yet another possibility potentially undermining the interpretation of iconicity effects in terms of stronger links between phonology and semantics: the effect of iconicity could be an effect of concreteness rather than iconicity *per se*. Although there are many examples of iconic signs for abstract concepts (often expressing concepts metaphorically, such as knowledge and mental states being located in the head, Taub, 2001), concrete concepts offer far more possibilities for iconic expression of perceptual and motoric attributes in sign forms. In spoken language it has been well established that concrete words have processing advantages over abstract words on a wide variety of tasks (e.g. James, 1975; Kroll & Merves, 1986; Paivio, Yuille & Smythe, 1966; Schwanenflugel & Stowe, 1989; Walker & Hulme, 1999), typically explained in terms of sensory-motor information associated with concrete concepts (e.g. Paivio, 1971, 1986, 1991, 2007; see Kousta et al., 2011; Newcombe, Campbell, Siakaluk & Pexman, 2012). Such accounts have a similar flavor to the explanations put forward about iconicity. Thompson et al. (2009, 2010) argued that iconic signs would activate their meanings faster, more automatically or more efficiently due to the transparent mappings between meaning and form; but if instead iconic signs do indeed tend to be more concrete than non-iconic signs, enhanced access to meaning could simply reflect the greater availability of sensory-motor information for the more concrete signs that also happen to be iconic. Thompson et al. (2010) attempted to mitigate this issue by taking ratings of concreteness and imageability (based on English translations of the BSL signs used) into account, but this was possible only to a limited extent as such ratings were only available for about half the BSL items in the study. Although there are still no concreteness norms available for BSL signs to date, we can obtain a rough estimate of the correspondence between iconicity and concreteness by using concreteness ratings for English translations of BSL signs, using a very large set of English norms (Brysbaert, Warriner & Kuperman, 2014). Of the 300 BSL signs for which Vinson, Cormier, Denmark, Schembri & Vigliocco (2008) collected iconicity ratings, concreteness ratings of English translations were available for 263. Iconicity and concreteness were highly correlated ($r(261) = .501$, $p < .0001$): barely any of the signs rated as high in iconicity were abstract (e.g., of the 50 most highly iconic signs, with iconicity ratings >5.6 on a 1–7 scale, only five had concreteness ratings less than 4 on Brysbaert et al.'s 1–5 scale). Such a strong relationship highlights the need to carefully consider the role of concreteness, in order to establish whether effects can be unambiguously attributed to iconicity.

To summarize, the evidence reviewed here suggests that iconicity in sign language may have task-specific effects: it facilitates semantic decisions and picture–sign matching, interferes with phonological decision and translation, and has no effect on priming in lexical decision. It also remains unclear how much of the previous effects attributed to iconicity may instead be due to the confounding factor of concreteness, as most highly iconic signs also tend to be highly concrete. It may therefore be premature to demand that language processing accounts built with the core assumption of arbitrariness need to account for effects of iconicity.

The present studies

If iconicity mediates between linguistic forms and bodily experience (Perniss et al., 2010), it must have reliable and consistent effects on language processing across tasks. In the present paper, we present a series of three studies that respond to specific criticisms raised to previous work and together provide a comprehensive test of iconicity effects in sign recognition and sign production. In Experiment 1 we use a picture–sign task in BSL to replicate the findings previously reported for ASL (Thompson et al., 2009) and, crucially, to assess whether iconicity effects are modulated by the typicality of the semantic features expressed iconically in the signs.

In Experiment 2 we explore further how iconicity affects phonological decisions by selecting a task based on a phonological property specifically linked to sign iconicity. In particular, we focused upon the relationship between movement of the hands and movement of referents (i.e., whether the physical motion in producing a sign includes upward or downward movement, for signs where movement upward or downward was more or less iconic of movement of the corresponding referent), and statistically controlling for concreteness. If the interfering effects of iconicity observed previously (Thompson et al., 2010) are related to the automatic activation of meaning, here we should find that iconicity facilitates, rather than interferes with responses. On the other hand, if iconicity effects in that study were due to other, uncontrolled differences between more and less iconic signs (such as morphemic status, Bosworth & Emmorey, 2010), or simply the consequence of the form-related focus of the task, iconic signs should show a disadvantage relative to non-iconic

signs, replicating the pattern observed by Thompson et al. (2010) for a different phonological parameter.

Finally, in Experiment 3 we turn to language production, using picture naming, a very natural task for which concerns about strategic, metacognitive or task-specific demands are vastly reduced. If iconicity has a general facilitatory effect linking meaning and form, a straightforward prediction is that iconicity should facilitate picture naming. If instead the link between meaning and form is arbitrary, iconicity should not play a role in picture naming once potential confounds like frequency, familiarity and concreteness are controlled, further suggesting that the apparent iconicity effects reported in comprehension and translation should instead be considered as more strategic or task-specific in nature.

Experiment 1

In the picture–sign and picture-word matching task conducted by Thompson et al. (2009), ASL signers were faster to respond when an iconic feature of the sign was made salient in the picture, while no such difference was observed in non-signers responding to English words or ASL signs. The semantic features evoked by the iconic properties of signs may be typical features of a concept (e.g., the BSL sign OWL which is iconic of an owl's eyes; see Fig. 2) and, thus, also likely to be represented when non-signers produce gestures referring to the same referent (e.g. Ortega & Morgan, 2010). Sign iconicity, however, can also highlight less typical features of meaning: those elements of meaning that do not seem salient for people who do not know a sign language. For example, the BSL sign TENT (see Fig. 2) is iconic of a tent's triangular shape, although the same sign is also used for tents with other shapes and, according to feature generation norms (participants' judgements of the features of meaning which are relevant to describing and defining a particular word, e.g. McRae, Cree, Seidenberg & McNorgan, 2005; Vinson & Vigliocco, 2008) it is not typical of the concept "tent".

We selected two different types of iconic signs for Experiment 1, varying in the extent to which the iconic property expresses a typical feature (high typicality or low typicality, according to English feature norms: McRae et al., 2005), and presenting salient vs non-salient pictures

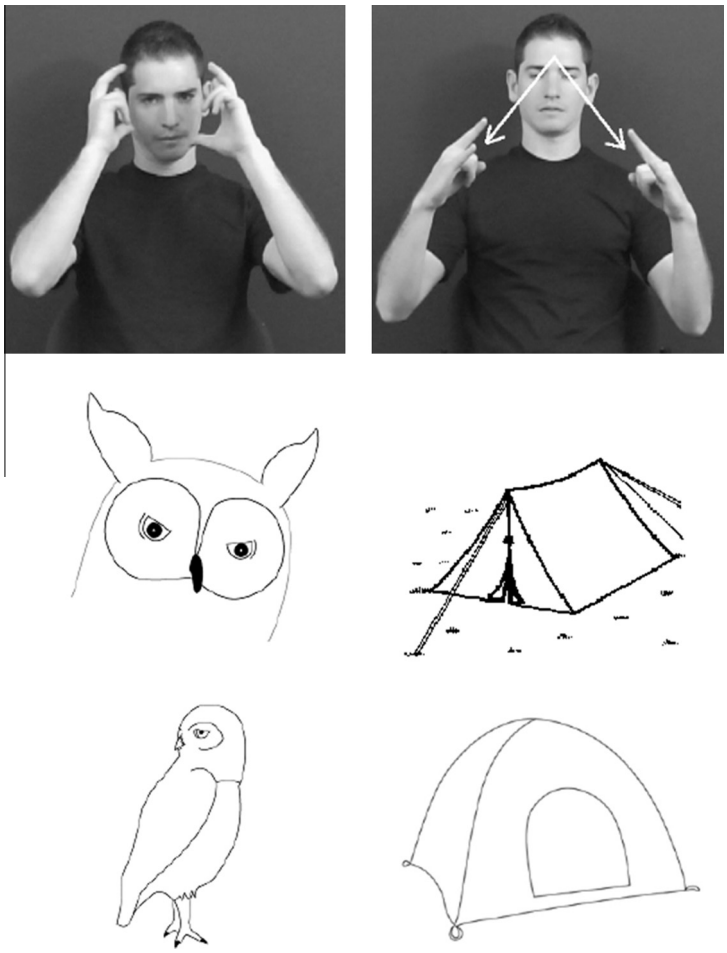

Fig. 2. Still image from BSL signs OWL (upper left) and TENT (upper right). Pictures reflecting the iconic property saliently appear in the middle panels; pictures in which the iconic property is not salient appear in the lower panels.

for each (as illustrated in Fig. 2). Like Thompson et al. (2009) we compared the performance of BSL signers to non-signers who saw video clips of a speaker producing English words (none of which are iconic). Our first prediction is that we should replicate the main findings of that study: a group (BSL signers vs. non-signers) × picture salience interaction such that salient pictures should facilitate responses for BSL signers, and not for non-signing English speakers (thus generalizing findings from ASL to BSL, a historically unrelated sign language). The crucial prediction concerns the interplay of iconicity and typicality. If the results from Thompson et al. (2009) did not arise due to iconicity *per se* but instead due to a confound related to typicality of the iconic features highlighted in pictures in that study, here we should see the same effect manifest instead as a group × feature typicality interaction. If instead iconicity is playing a role, the group × picture salience interaction should instead be observed even when typicality is controlled. Finally, a significant three-way interaction (group × picture salience × typicality) would be observed if iconicity only strengthens the link between meaning and form when a less typical feature is represented in the sign.

Method

Participants

Sixteen deaf BSL signers (9 women; 7 men, age range 18–51, median = 24) were recruited primarily through personal contacts in Birmingham, Preston and the greater London area. Twelve were native or early BSL signers (eight native, the other four exposed to BSL before age 5); the remaining four were exposed to BSL later (between ages 6–12, all in residential deaf schools). All use BSL as their preferred and primary language. Nineteen English speakers (10 women, 9 men; age range 18–37, median = 24) were recruited through the participant pool at the UCL Department of Psychology. They all reported no knowledge of BSL beyond the manual alphabet (a set of signs representing the English alphabet, used to fingerspell English words).

Materials

In order to select materials for which picture salience and typicality could be manipulated, we went through an extensive selection process described in Appendix A. Once this was completed, we were left with a set of 54 signs/words, each with two pictures (one in which the iconic property in BSL was expressed saliently, and one in which it was not), with >80% name agreement in both languages (i.e. 80% or more people produced the same sign or word, given a target picture). Of these, 27 items were of high typicality according to English norms, and 27 were of low typicality. Ultimately eight of these were excluded (see below), leaving 25 high typicality and 21 low typicality items in the analyzed set (see Appendix B for a list of items and their characteristics). A total of 188 filler trials were included (40 in which the picture matched the sign/word, and 148 in which the picture and sign/word did not match), so that there were an equal number of match and mismatch trials in the experiment as a whole; see Appendix A for the details of how filler items were prepared. Finally, we added additional signs/words and pictures not otherwise used in the experiment to be used in practice trials, including some iconic and some noniconic signs, also with an equal number of match and mismatch responses.

The video clips of BSL signs used in the experiment were produced by a deaf native BSL signer. For signs with more than one variant, we used the sign form for which we had obtained additional normative data (e.g. Vinson et al., 2008). Each sign began and ended with the hands in a rest position and thus all began with a preparatory/-transitional phase in which the signer's hands moved into the starting position. English words were also video recorded and produced by a native speaker of British English. Each video clip was converted to AVI format (25 frames/s, 720 × 576 pixels on a 1024 × 768 display).

Design

For analysis, we only considered the crucial experimental trials (in which the picture matched the sign or word). Our first variable of interest was Picture salience (salient: BSL iconic feature is expressed saliently in the picture, vs. non-salient). This was factorially combined with Typicality (typical: iconic feature is typical according to English norms), participant Group (deaf BSL signers; English-speaking non-signers), and Order of presentation (first block vs second block, in case of repetition effects); thus considering a fully factorial design up to and including four-way interactions. We then analyzed the effects of these factors on decision response latencies, employing mixed-effects models with crossed random effects for participants and items, using restricted maximum likelihood estimation.[2] To do so we used the package lme4 (version 1.0-4: Bates, Bolker, Maechler & Walker, 2013) running in R version 3.0.1 (R Core Team, 2013). Before fitting the models we contrast-coded all of the fixed effects (in all cases, they were factors with two levels) to ensure that interaction terms were orthogonal to main effects. In addition to random intercepts for participants and items we also included random slopes in the model: by participants, Picture salience, Typicality, and Order of presentation; by items, Picture salience, Group and Order of presentation thus starting with a maximal, "design-driven" random effects structure (Barr, Levy, Scheepers & Tily, 2013).

Procedure

Stimulus presentation was carried out using DMDX version 3.2.2.3 (Forster & Forster, 2003). For experimental items, order of presentation was balanced across participants and items (e.g., a salient picture would occur before a non-salient picture on half the trials, and vice versa, and this order would be reversed for the next participant). Filler items were also assigned in advance to either the first or second block of the experiment to ensure that repetition

[2] We originally conducted two separate factorial ANOVA, one with subjects as random effect and the other with items (F_1/F_2 analysis), but later adapted our approach for the sake of consistency among the studies reported here. The main findings from this experiment do not change if F_1/F_2 ANOVA is conducted instead.

Table 1
Trimmed correct response times as a function of language × typicality × salience. Values in brackets indicate lower and upper quartiles (50% of predicted data points in that cell fall within this range).

	BSL		English	
	Non-salient picture	Salient picture	Non-salient picture	Salient picture
Low typicality	773 (640, 883)	756 (622, 869)	900 (761, 1022)	913 (770, 1037)
High typicality	792 (660, 899)	761 (631, 867)	902 (757, 1026)	902 (764, 1019)

of some filler items was comparable to the repetition of experimental items. Order of presentation was random within each block.

Participants were first given instructions in BSL or English, told that the task involved sign or word recognition: that they would see a picture followed by a sign or word, and their task was to press one key ("j" on a standard UK QWERTY keyboard) if the sign/word matched the picture and another ("f") if it did not match. They were asked to respond as quickly as possible while still maintaining accuracy. The experiment began with 20 practice trials. Each trial began with a fixation cross displayed for 1500 ms, followed by the picture which was displayed for 800 ms, followed by a video clip of the target sign or word after which a blank screen appeared. Responses were permitted from the start of the video clip, until a 2500 ms timeout. Response times were measured from video onset, and accuracy was recorded automatically. A blank screen was displayed for 1000 ms before the next trial started. Breaks were given every 27 trials.

Once the experiment was finished, we presented the BSL participants with each of the iconic pictures for target items and asked them to produce their sign. Cases of variation were taken into account in the analyses (see Section 'Results').

Results

We first excluded trials where the picture did not match the sign/word as well as all filler trials. After doing so we then excluded trials for BSL signers where a person's preferred sign did not express our intended iconic feature (although we permitted minor variations in sign form which expressed the same iconic feature, as excluding such cases did not affect our results). Five items were excluded from all analyses due to extensive variation among participants' preferred signs and/or low response accuracy rate; three more items were removed from the set because the salient and nonsalient pictures were not sufficiently different (based on ratings by non-signers who judged the extent to which each picture visually matched the video clip of the BSL sign). This left 48 items for analysis. We then excluded four participants (two BSL signers, two non-signers) whose overall accuracy on the task was less than 80%.

Analyses of response latencies excluded all errors (3.7%), as well as trials more than three standard deviations from a participant's mean (1.5%). We then conducted an analysis including the 2 × 2 × 2 × 2 full factorial design as described in the Design section (full random effects structure for participants and items). As this model did not converge, we removed the random slopes of the control variable Block by participant and item, and the variable Picture salience by item, before proceeding further (see Barr et al., 2013).

Of the main effects, only Group and Block approached conventional levels of significance. There was a tendency for BSL participants to respond faster than English speakers (β (BSL–English) = −134, 95% confidence interval (CI) for the parameter estimate (−269, 2), $t = -1.935$), which might be attributed to earlier recognition points for signs vs. speech (see Emmorey & Corina, 1990; Grosjean, 1981). There was also a tendency for responses in the second block to be faster than the first (Block: β (First–Second) = 31.1, 95% CI (−1.5, 63.7), $t = 1.873$); for main effects of Salience and Typicality $|t| < 1$).

The crucial interactions concern the factors of Typicality, Salience and Group (see Table 1). Of these, only the Group × Salience interaction was reliable ($\beta = -29.9$, 95% CI (−56.9, −2.9), $t = -2.171$): the estimated parameter β reflects the BSL–English difference in the effect of salience. The other interactions involving these factors were not reliable (all $|t| < 1$); None of the other interactions were reliable (Group × Block $\beta = -59$, 95% CI (−124, 6), $t = -1.781$; Group × Block × Salience $\beta = 47$, 95% CI (−12, 105), $t = 1.567$; Salience × Typicality × Block $\beta = -78$, 95% CI (−163, 6), $t = -1.814$; all other $|t| < 1$). To better understand the interaction between Group and Salience, we fit separate models for BSL and for English, including only the main effects of Salience, Block and Typicality in each one. In BSL the effect of Salience was reliable ($\beta = -24$, 95% CI (−43, −5), $t = -2.51$) but it was not in English ($\beta = 5.5$, 95% CI (−14, 25), $t = 0.554$). Pictures in which the iconic BSL feature was expressed saliently (like OWL with prominent eyes as in Fig. 2) yielded faster responses than pictures in which it was not (OWL with the eyes less prominently featured), but only in BSL and not in English (see Fig. 3).

As feature typicality did not predict response times, we conducted one additional analysis to see whether feature distinctiveness, the extent to which an iconic feature is shared among few concepts, might have played a role instead. After all, many iconic signs involve features that distinguish a concept from its nearest neighbors (e.g. trunk for ELEPHANT, beard for GOAT, popping-up motion for TOASTER, flashing light for AMBULANCE). Following the same logic as for typicality, if iconicity effects are item specific, for example restricted to very distinctive features, or if this task exaggerates effects when a less distinctive feature is highlighted, we should find that picture salience interacts with distinctiveness. To test this we fit a model to BSL signers' data including only Picture Salience × Distinctiveness.

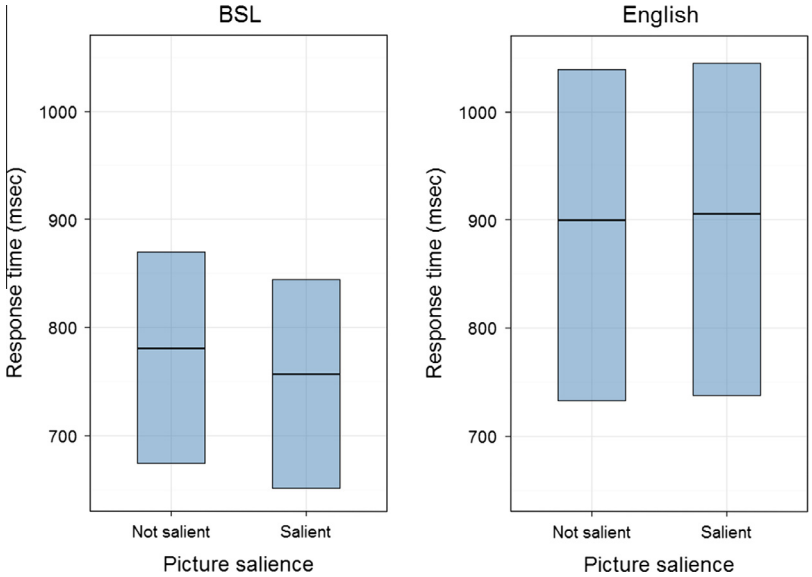

Fig. 3. Trimmed correct response times as a function of language and picture salience. BSL signers were faster when the pictures made an iconic feature of the sign salient; English speakers were not. Solid lines represent estimate of the mean; shaded areas depict upper and lower quartiles of predicted RTs in each condition (50% of predicted values fall within this range).

We operationalised distinctiveness for iconic features of the BSL signs following McRae et al. (2005): 1/number of concepts for which a feature is listed (in McRae et al.'s set of 541 concrete noun concepts), regardless of that feature's weight (typicality). A value of 1 thus indicates that a concept is unique among this set in having that feature listed by McRae et al.'s participants, and lower values indicate that a feature is shared among more concepts. For some concepts the iconic feature was not listed at all ($n = 13$) and these items were left out of this analysis. Neither the main effect of Distinctiveness nor its interaction with Picture salience was reliable (both $|t| < 1$) while the effect of Picture salience remained ($\beta = -40$, 95% CI (-81, -0.3), $t = -1.98$). Finally, we carried out one last set of analyses to see whether the type of property expressed by iconicity (perceptual, motor, or both perceptual and motor iconicity) modulated the effect of Picture salience in BSL signers. This was not the case (picture salience $\times$ iconicity type $|t| < 1.2$).

Discussion

In Experiment 1 we replicated the results of Thompson et al. (2009) in a different, historically unrelated, signed language: pictures which saliently express an iconic property of a sign facilitate BSL signers' responses to that sign, compared to pictures which do not express iconicity saliently. Importantly, no such difference is observed for English speakers. Crucially, this effect does not differ when comparing signs whose iconic phonological properties reflect non-linguistic features that are typical to non-signers, with signs whose iconic properties are less typical to non-signers. If the previous findings had been due to a confound of the typicality of the features highlighted in the pictures, we should have found differences between typical and atypical features. Because interactions involving picture salience and feature typicality were not observed, nor were there interactions with feature distinctiveness or with the type of iconic property in the final analyses we reported, the most straightforward interpretation is that iconicity has a general facilitatory role, reflecting an overall enhancement of the accessibility of non-linguistic features that are linked to iconic phonological features (at least, where properties of concrete objects are concerned).

However, these results must still be interpreted with some caution. Although the items used in our experiment (and by Thompson et al., 2009) were embedded in a large number of fillers for which feature salience was not manipulated, it remains possible that the particular demands of this matching task could be responsible for the effects of iconicity (as suggested by Baus et al., 2012; Bosworth & Emmorey, 2010; Emmorey, 2014). The crucial manipulation of feature salience across picture pairs (e.g. an owl with the eyes very prominent, vs an owl with less prominent eyes as pictured in Fig. 2) may have drawn signers' attention to this manipulation and its relationship to the form of the sign itself. If signers are paying more attention to specific features of the pictures, rather than the main referent illustrated by the features, this could facilitate responses to iconic signs either by activation of the name of an iconic feature rather than the sign itself (i.e. phonological correspondence between the signs EYES and OWL), or due to visual imagery of the feature and its overlap with the form of the subsequently presented sign as a means of structure mapping (Emmorey, 2014). In English this would not be the case as there is no form similarity or imagistic overlap between features and the subsequently presented words.

Finally, one may wonder whether the verbal feature norms for English words adequately reflect the properties

of concepts that might be iconically represented in BSL. The items listed in Appendix B suggest that iconic features classified as low typicality on the basis of features might not truly be less typical, but may instead be difficult to express by participants in a feature generation task. For example, the manner in which a pipe is held, or a piano is played, or a bicycle is ridden seems essential to the way these items are used, but they are relatively difficult to express as a single written feature. The same is true of many shapes that may be highly distinctive but hard to express verbally (see Vinson, 2009). As the same issue also applies to measures of distinctiveness derived from these measures, from this study we cannot entirely reject the possibility that iconicity effects may be restricted to a subset of signs.

As a result, in order to develop accounts of how iconicity affects lexical processing it is necessary to employ converging methods, moving away from tasks involving explicit semantic decisions or comparison processes and instead employing tasks that are not so subject to attentional or strategic criticisms. In Experiment 2 we address this by investigating how iconicity affects the processing of single signs, using a phonological decision task where facilitatory effects are predicted if iconicity affects online processing.

Experiment 2

The results of Experiment 1 indicate that the meaning-form correspondence for iconic signs facilitates comprehension above and beyond what is expected on the basis of semantic feature typicality, but as previously discussed, these results could also be explained in terms of task-specific demands. Experiment 2 responds to this criticism with a task that focuses on phonological properties of signs. Moreover, this experiment directly tests the account put forward by Thompson et al. (2010) according to which interference was found in their phonological-decision task because semantics was activated by the iconic properties of the sign, but being irrelevant to the decision, hindered it. This account predicts that if the iconicity of the sign is task-relevant (movement features that map onto real-world movement) then it should facilitate the decision. Meaningful movement is seen, for example, in the BSL sign ROCKET, produced with the index finger extended upward in front of the body and a long upward movement. The phonological movement of the sign directly represents the real-world movement of rockets. Such iconic movement encoded in a sign tends to indicate the way something is handled and used, or represent how an object or action moves through space (Schembri, 2003). At the other end of this continuum are signs with upward/downward phonological movement that do not map onto real-world movement. The sign HOUSE, for example, has a downward movement indicating the outlined shape of a house, not its movement (see Fig. 4). In Experiment 2, signers made a phonological decision (does the sign move up or down?).

However, if the iconicity effects we reported in Thompson et al (2010) were due to other confounding factors, either we should replicate the interference effect here, or perhaps we should fail to observe any iconicity effect once we control for variation in concreteness and other factors correlated to iconicity.

Method

Participants

Twenty-eight BSL signers participated in this study (16 deaf, 12 hearing), all recruited from Deaf communities in London, Birmingham and Edinburgh. Of the deaf BSL signers (8 women, 8 men, age range 21–66, median = 31), four were native signers, three were early signers (exposed to BSL before age 5), and the rest were late signers (exposed to BSL between age 6 and 25). Of the hearing signers (7 women, 5 men, age range 21–62, median = 31.5), four were native signers and the rest were late (learning BSL between age 12 and 35). All the BSL signers had been signing for a minimum of three years at the time of testing). We further recruited 14 hearing non-signing monolingual English speakers with no previous knowledge of BSL other than the fingerspelled alphabet (six women, eight men; age range 20–69, median age 26).

Materials

We started with 170 video clips of BSL lexical signs with clear upward or downward movement, for which we collected norms for familiarity, motion iconicity and general iconicity from ten BSL signers (7 = native, 3 = late, learning BSL after age 5). Motion iconicity ratings are a measure of perceived upwards or downwards movement for the real world object/action of any given sign with phonological movement up or down (i.e., how well the up/down meaning of a sign maps onto its phonological movement); see Appendix C for details of these norms. From these norms we selected a final set of 108 signs (plus 16 practice items), of which 54 have a single upward motion and 54 a single downward motion. Items in upward and downward motion categories were balanced for iconicity, familiarity, concrete or abstract meaning, grammatical class and motion iconicity. Video clips of all signs were created with a native deaf signer as the model.

Design

Our main variable of interest was motion iconicity, treated as a continuous measure. We assessed its effects on BSL signers' response times taking into account other potentially confounding variables, and in contrast to nonsigners who do not have linguistic experience with the same materials. Non-signers could participate in the same experiment because the movement decision required no understanding of sign meanings; testing the group × iconicity interaction allows us to rule out the possibility that iconicity effects are due to other characteristics that might covary with iconicity (e.g., differences in signs' onset or duration, or varying perceptual difficulty of signs with different features). We employed mixed-effects models with crossed random effects for participants and items, using restricted maximum likelihood estimation (as in Experiment 1 using lme4 version 1.0-4, Bates et al., 2013; and R version 3.0.1, R Core Team, 2013). Predictors included motion iconicity ratings, familiarity ratings (by BSL signers; Vinson et al., 2008)

Fig. 4. Still images from BSL signs that move upwards (upper panels) and downwards (lower panels). Signs on the left are rated low on motion iconicity; signs on the right are rated high on motion iconicity. Upper left: BOTTLE (moves upward to reflect the shape of the bottle but does not refer to motion). Upper right: ROCKET (movement upward corresponds to a rocket taking off). Lower left: MAN (movement downward may correspond to shape of a beard but does not refer to motion). Lower right: RAIN (movement downward corresponds to rainfall).

and direction of sign movement (up or down). This factor is particularly important to take into account as all of the video clips started with initial upward motion (the signer's hands always began from a resting position in his lap); as a result, all signs moving downward would have started with upward motion to reach their starting points. Initial models included group (signer, nonsigner) and all possible two-way interactions involving group; non-significant interactions were dropped and "final" models fit without them. The models had random effects structure including random intercepts, motion iconicity and motion direction slopes for participants (nested in hearing status and age of learning BSL) and random intercepts for items.[3]

In a separate set of models, we included concreteness ratings for the English translations of the signs (Brysbaert et al., 2014)[4], in order to make sure that any effects of iconicity could not just be attributed to differences in concreteness as the two are significantly correlated in this set (motion iconicity – English concreteness $r(93) = .416$, $p < .001$). Finally, we fit a set of models, using data from BSL signers only and testing whether effects of motion iconicity and other factors depended on participants' hearing status (deaf or hearing) or age of learning BSL (early or late).

Before fitting the models, we centered all continuous predictors (motion iconicity ratings, familiarity ratings, nonsigner response times) and contrast coded all categorical predictors to ensure that interaction terms were orthogonal to main effects. In a first series of models we considered not only linear but also nonlinear (quadratic) transformations of continuous measures in initial model fits, with nonlinear transformations retained only when they provided significantly better fit over linear terms (determined via likelihood ratio tests).

Procedure

After giving their consent to participate, signing participants were presented with video-recorded instructions by a native BSL signer. Participants were instructed to decide as quickly and accurately as possible if a sign moves upwards or downwards and respond by pressing a computer key. The two keys used, D and K, were

[3] We also fit models including random slopes by item for participant group, deafness and direction of movement, but as these more complex models did not converge, and these factors were not of central importance to the questions we are addressing, they were not retained in the analyses we report.

[4] In this analysis we excluded the 10 BSL signs that did not translate into a single word in English, or did not appear in Brysbaert et al.'s (2014) ratings.

counterbalanced across participants for up/down responses. Participants were given practice trials to begin with and invited to ask clarification questions before beginning the actual experiment. During the experiment participants were given frequent opportunities to take breaks. Stimuli were presented on a PC, using DMDX v.3.2.2.3. Each trial began with a fixation cross (displayed for 400 ms), followed by a single BSL sign (720 × 576 pixels). Participants were able to respond as soon as the clip began to play. Once a response was made, or after a 3000 ms timeout, there was a 250 ms delay before the fixation cross appeared again for the next trial.

Results

We first excluded all items with accuracy rates below 75% ($N = 4$), leaving 104 in the data set (average 93.6% correct). We then excluded all error trials and those trials with RTs more than three standard deviations from a participant's average. Average trimmed correct response latencies in this task were 1166 ms ($SD = 270$). We then turned to the data from the English-speaking non-signing participants, considering the 104 items that were kept for analysis based on the signers' responses. With the exception of one participant who responded at chance levels and was excluded, non-signers were highly accurate at the task (average accuracy, 90.9%), and their response times were comparable to those of the signers (average trimmed correct response latencies = 1219 ms, $SD = 377$).

We first considered whether it was necessary to include nonlinear transformations of the continuous predictors (motion iconicity and familiarity ratings); likelihood ratio tests indicated no significant benefit of adding nonlinear transformations to the models. We then fit additional sets of models varying in the inclusion of interactions (all 2-way interactions involving group, plus main effects). The interaction between group and familiarity was not reliable: a model including this interaction did not offer sufficient improvement over the model without it ($\chi^2(1) = 0.091$, $p = .762$; log-likelihood ratios both = −27,874) and so was not included; the other two interactions were retained.

In this model, the main effect of familiarity was reliable ($\beta = -28$, 95% CI (−43, −12), $t = -3.44$), showing that relatively substantial amount of the variance in response latencies is captured by familiarity of the signs as rated by people who know BSL. The main effect of motion direction was not significant (β of the difference (Upward–Downward) = −19, 95% CI (−57, 20), $t = -0.96$), nor was the effect of group (β of the difference (Signer–Nonsigner) = −4, 95% CI (−96, 89), $t = -0.08$). The main effect of motion iconicity did not reach significance ($\beta = -13$, 95% CI (−27, 1.1), $t = -1.81$). However, these main effects were qualified by significant interactions involving group (motion direction × group interaction $\beta = 53$, 95% CI (13, 105), $t = 2.53$); motion iconicity × group interaction $\beta = -20$, 95% CI (−31, −9), $t = -3.44$).

To better understand these interactions, we fit separate models for signers and nonsigners including main effects of familiarity, motion direction and motion iconicity in each. For signers there were main effects of familiarity and motion iconicity only (familiarity $\beta = -27$, 95% CI (−43, −11), $t = -3.36$; direction of motion $\beta = 12$, 95% CI (−22, 45), $t = 0.68$; motion iconicity $\beta = -23$, 95% CI (−37, −9), $t = -3.28$). For nonsigners, there was also a main effect of BSL familiarity, but in contrast to signers there was a positive effect of motion direction and no effect of motion iconicity (familiarity $\beta = -29$, 95% CI (−48, −9), $t = -2.929$; direction of motion $\beta = -49$, 95% CI (−90, −9); motion iconicity $\beta = -3$, 95% CI (−20, 15), $t = -0.321$. Signs moving upward tended to elicit faster responses for nonsigners only. Crucially, we found effect of motion iconicity only for signers (see Fig. 5 for plots of the results).

The effect of motion iconicity was not driven by differences in concreteness between more and less iconic signs; concreteness was not a significant predictor in the models that included it, while the motion iconicity effect remained reliable (e.g. in the model including motion direction × group and motion iconicity × group , concreteness $\beta = 16$, 95% CI (−1, 32), $t = 1.89$; motion iconicity × group $\beta = -22$, 95% CI (−34, −10), $t = -3.49$). We then fit additional series of models on the signers' data only, fitting separate sets of models that included hearing status (Deaf–Hearing) and age of exposure to BSL (Early–Late) along with familiarity, direction of motion and motion iconicity. Hearing status did not exert any effects on response time (non-significant main effect and two-way interactions involving hearing status: all $|t| < 1.5$), and the same was true of age of exposure to BSL (all $|t| < 1.4$). Crucially in both of these models, the effects of familiarity and motion iconicity remained reliable ($t < -3.2$ in all cases): independent of all other factors tested, signs with high motion iconicity ratings tended to elicit faster responses than signs with lower motion iconicity (see Fig. 5).

A somewhat surprising result is the main effect of familiarity. Why would familiarity of BSL signs, as rated by people who are fluent in BSL, predict performance of nonsigners just as well as signers? We considered the possibility that this effect may reflecting some more general correlate of familiarity in one final analysis. Based on our observation that familiarity is correlated with duration of the video clips ($r = -.366$, $p < .001$) we fit a model that included duration (in video frames) along with familiarity, motion iconicity, direction of motion and the main effect and interactions involving group. Indeed, duration in frames was a reliable predictor ($\beta = 12$, 95% CI (9, 15), $t = 7.15$) with faster responses to shorter videos, while the effect of familiarity was no longer significant ($\beta = -12$, 95% CI (−25, 2), $t = -1.72$). Crucially, the interactions involving group remained significant in this model (motion direction × group $\beta = 59$, 95% CI (15, 104), $t = 2.59$; motion iconicity × group $\beta = -20$, 95% CI (−31, −10), $t = -3.27$).

Discussion

In Experiment 2 we found that iconicity facilitated signers' decisions about form-related characteristics of BSL signs (whether the sign moves up or down), consistent only with the prediction derived from Perniss et al. (2010). This finding complements Thompson et al. (2010) where response times were slowed when participants

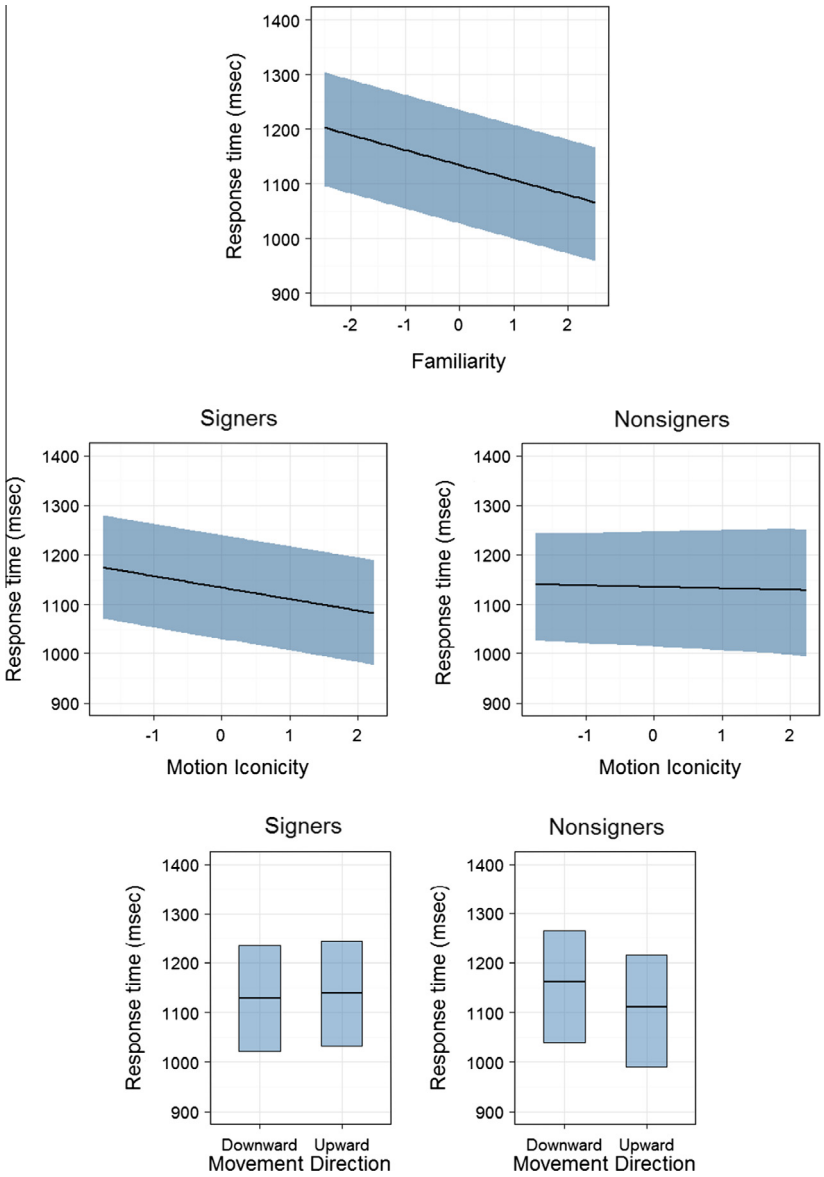

Fig. 5. Partial effects of significant predictors in analysis of correct RTs in Experiment 2. Upper: main effect of familiarity: responses by signers and nonsigners alike are faster for more familiar signs. Middle: motion iconicity × group: signers make faster responses for more iconic signs; nonsigners do not. Lower: movement direction × group: non-signers make faster responses for signs that move up; signers do not. Continuous predictors (centered for analysis): solid lines indicate estimate of the mean; shaded areas depict upper and lower quartiles (50% of the predicted RTs occur within this range). For categorical predictor (movement direction) the shaded area depicts upper and lower quartiles of predicted RTs in each condition.

made a phonological decision not related to iconically-encoded meaning (whether handshape was straight or curved). In that study, automatic activation of meaning by iconic properties interfered with the form-related decision. Here, instead, because items were selected so that their iconic properties related to the phonological decision, we found facilitation. Because these effects were observed only for signers and not for people who do not know BSL, we can further conclude that they depend on language experience and are not simply related to visual or temporal properties of the signs (which would have affected signers and nonsigners alike, as seen in the effect of video duration). The fact that we could change the effect of iconicity from an interference effect to a facilitatory effect changing whether the iconic property was relevant to the task, further renders a strategic account of iconicity effects less plausible as both experiments (Thompson et al., 2010 and Experiment 2) used a phonological decision task.

These results converge in indicating that effects of iconicity arise due to links between properties of form and components of meaning: the conceptual features derived from action or perception that are represented in lexical forms. Making an iconic feature more salient in a picture, as in Experiment 1 (and Thompson et al., 2009)

would enhance recognition of that specific sign as a consequence of additional activation of one or more of its conceptual features. But in phonological decisions, the conceptual features activated by an iconic sign are likely to be shared across multiple lexical items (e.g., many signs include iconic upward phonological movement corresponding to meaningful motion upwards). Because the task does not require participants to recognize or select a specific lexical item, only to respond as soon as they become aware of the form distinction, mutual activation of signs sharing a feature like "upward movement" can facilitate responses as we observed in Experiment 2. More general activation not tied to specific features of perception and action seems less likely given the findings from Bosworth and Emmorey (2010) that semantically related concepts with different iconic signs (reflecting different features) do not exhibit increased semantic priming compared to semantically related concepts that do not both have iconic signs. But it still remains difficult to rule out the possibility that structure mapping could be responsible for these findings as well (see Emmorey, 2014); lack of alignment between motion features in noniconic signs and direction of motion would result in facilitation for iconic signs in the task employed in Experiment 2.

For a stronger test of the role of iconicity, we need to switch from receptive language tasks involving perception and/or comprehension of iconic signs. If iconicity has a general role in language processing, it should also affect sign production, a domain that has been almost entirely neglected in studies of iconicity to date (with the exception of sign repetition tasks, e.g. Ortega & Morgan, 2010). Experiment 3 employs picture naming, a highly natural and well-practiced task which avoids concerns about strategic or task-specific effects that could apply to many previous studies of iconicity. Moreover, picture naming should not involve structural alignment (Emmorey, 2014) as there is no comparison, simply semantically driven lexical retrieval and production. If iconicity affects processing due to stronger links between features of meaning and iconic forms, picture naming should still be faster for iconic signs than non-iconic signs even after controlling for other factors; while no effect of iconicity should be observed in naming if the underlying lexical representations are arbitrary, or if previous effects of iconicity are due to degree of alignment permitted in matching tasks.

Experiment 3

> "To name a picture can be considered an elementary process in the use of language".
>
> [Glaser, 1992, p. 61]

Picture naming has been used from the earliest days of experimental psychology research, in order to illuminate aspects of the fast and highly practiced processes by which words are produced (e.g. Brown, 1915; Cattell, 1886; Hawthorne, 1934; Stroop, 1935), and data from picture naming has served as a crucial basis for theories of cognitive representation (e.g. Glaser, 1992; Johnson, Paivio & Clark, 1996) and language processing (e.g. Levelt, 1989; Levelt et al., 1999). It is widely accepted that picture naming occurs via the following general processes: object recognition, conceptual processing, retrieval of an abstract lexical representation, retrieval of form and then articulation (e.g. Butterworth, 1989; Dell, 1986; Dell, Schwartz, Martin, Saffran & Gagnon, 1997; Garrett, 1980, 1984; Levelt, 1989; Levelt et al., 1999; MacKay, 1987; Rapp & Goldrick, 2000; Roelofs, 1993). These views also have in common the assumption of arbitrariness and all maintain that lexical retrieval entails two separate steps: semantic and phonological retrieval, although they differ widely on the structure and flow of information between levels (Vigliocco & Hartsuiker, 2002).

In spoken languages, picture naming latencies are affected by properties of the words to be named. Early work showed that frequency of occurrence predicted naming latencies: more frequent words being produced faster than less frequent words (Oldfield & Wingfield, 1964), with subsequent studies also suggesting an independent role of age of acquisition (Carroll & White, 1973) and more recently Belke, Brysbaert, Meyer & Ghyselinck, 2005; Cortese & Khanna, 2007; Hernandez & Fiebach, 2006). Although both these factors are correlated and both facilitate naming (more frequent words and earlier-acquired words are named faster and more accurately), characteristics of speech errors in anomia suggest that AoA effects are particularly relevant for form retrieval, while frequency effects are relevant for both meaning and form retrieval as considered in two-stage models of lexical retrieval. In particular, Kittredge, Dell, Verkuilen and Schwartz (2008) found that frequency predicted error occurrence of multiple types (semantic phonological and omission errors; with stronger effects at the phonological level) while AoA only predicted phonological and omission errors, not semantic errors. Effects of AoA were also observed by Navarrete, Scaltritti, Mulatti and Peressotti (2013) in a delayed picture naming task further suggesting that AoA affects phonological retrieval. If similar two-stage architectures underlie sign and speech production, one would expect the effects of these variables to be comparable in BSL.

It is important to note that, as far as we are aware, no previous study has discussed how iconic wordforms would be retrieved. However, if iconicity has a general role in language processing, we should find effects in production as well as in comprehension. In production, if iconicity renders the links between semantic features and phonological properties stronger, it should aid the retrieval process. If this is the case, we may also expect that iconicity effects might be modulated by those same properties that also facilitate form retrieval in naming: frequency and AoA. Similarly, if iconicity facilitates form retrieval in production, it is also possible that age of exposure to BSL modulates iconicity effects in production. After all, late exposure (particularly as a first language) is correlated with poorer performance in many language tasks (e.g. Cormier, Schembri, Vinson & Orfanidou, 2012; Ferjan Ramirez, Leonard, et al., 2013; Ferjan Ramirez, Lieberman, et al., 2013; Mayberry & Eichen, 1991; Mayberry & Tuchman, 1985; Meier, 1991). If so, one might expect effects of iconicity to be greater for late learners in spite of the lack of such differences in comprehension. Therefore in Experiment 3 we tested the effect of iconicity

on picture naming, additionally considering its interaction with factors including frequency, AoA and participants' age of exposure to BSL.

Method

Participants

A total of 24 deaf adults (13 women, 11 men, age range 19–61, median 30.5) participated. One participant was eliminated because of technical problems, leaving 23 participants with analyzable data. Of these, 11 were native signers, four early (learning BSL before age 5) and eight were late signers, learning BSL at ages ranging from 6 to 24 (all had been signing for at least four years at the time of testing). All of them rated their BSL skills as 5 or higher on a 1–7 scale where 7 indicates nativelike fluency.

Materials

Experiment 3 was conducted as the pretest phase of an eye-tracking study (Thompson, Vinson, Fox & Vigliocco, 2013) used to familiarize participants with the pictures in that study. Participants produced signs for 424 black and white line drawings taken from various sources (Snodgrass & Vanderwart, 1980; Szekely et al., 2004; Thompson et al., 2009). Of those 424 signs produced, we analyzed 92 for which there were available norms for familiarity and iconicity (Vinson, Cormier, Denmark, Schembri & Vigliocco, 2008).

Design

Our main variable of interest was iconicity, treated as a continuous measure. We assessed its effects on BSL signers' naming latencies taking other potentially confounding variables into account, using the same type of analyses as we used Experiments 1 and 2: mixed-effects models with crossed random effects for participants and items, using restricted maximum likelihood estimation. Our predictors included Iconicity ratings and Familiarity ratings from Vinson et al. (2008), and Frequency of occurrence (log(count + 1) from Fenlon, Schembri, Rentelis, Vinson & Cormier, 2014), participants' Age of learning BSL (early or late) and Order of picture presentation. To account for response time differences that might arise due to differences in phonological complexity, we entered into the model a sign complexity rating based on Mann et al. (2010). This measure was obtained by adding up measures of complexity across different phonological parameters as follows: (1) Handshape: 0: signs which used one of 4 unmarked handshapes, ('B', '5', 'G' and 'A'; see Sutton-Spence & Woll, 1999), +1: all other handshapes and +1: any handshape change (e.g., from an 'A' to a '5'). (2) Movement: 0: signs with one movement, either internal movement (e.g., opening, closing, orientation change, wiggling) or path movement (e.g., straight, arc), +1: signs with both an internal movement and path movement, or more than one path movement. (3) Location: 0: signs produced in neutral space in front of a signer's body, +1: all other locations on the body, and +1: location change (more than one location in a given sign). (4) Hands: 0: one-handed signs, +1: two-handed sign, +1: two different handshapes (for two-handed signs only). The sum of these phonological complexity measures was entered into the model for each item.[5] All two-way interactions were included in initial models, removed if they were not warranted (based on model comparisons with and without a given interaction).

Random effects included intercepts by participant (nested in signer group) and item, as well as random slopes (by participants: Iconicity Familiarity, Frequency, Order; by items: Age of learning BSL and Order). We also took Age of acquisition (AoA) ratings for the BSL signs into account in secondary analyses (necessary because this measure is only available for 77 of the items included in the main analysis). As in Experiment 2 we also included Concreteness ratings for the English picture names (Brysbaert et al., 2014) in these secondary models. As with the main analyses these predictors were also included as random slopes by participants.

Before fitting the models we centered all continuous predictors and contrast coded all categorical predictors to ensure that interaction terms were orthogonal to main effects. In a first series of models we considered not only linear but also polynomial transformations of continuous measures in initial model fits, with nonlinear transformations retained only when they provided significantly better fit over linear terms (using likelihood ratio tests). We then tested models to consider whether iconicity interacted with other factors; including all two-way interactions involving iconicity.

Procedure

Participants were instructed to press the space bar, using their dominant signing hand and release it only when they were ready to produce the sign for the picture on the screen. Response times were measured from keyboard release. After production of a sign, the next trial began when the key was depressed again, allowing for self-paced breaks throughout. All sessions were video recorded.

Results

For all signs analyzed, we first examined video recordings and excluded productions of signs that differed from our available norms (due to dialectal variation), as well as sign errors (particularly productions which began with hesitations or stalls). We then excluded all those trials with RTs more than three standard deviations from a participant's average. This left an average of 21 tokens for each sign analyzed, with average trimmed correct response latencies of 1137 ms ($SD = 370$).

Among the items in the study there were 15 for which AoA ratings are not available, so we did not take AoA into account in the first analyses. As the models failed to converge with the maximal random effects structure, we removed the random slopes for the control variable Order of presentation (which also exhibited near-zero variance among random slopes in the maximal model) but retained it as a fixed effect. We then considered a model also including nonlinear (quadratic) transformations of

[5] We also fit a model of phonological complexity in which the different parameters were independent.

each continuous predictor to allow for the possibility that their effects might not be strictly monotonic; only order of presentation warranted inclusion of a nonlinear term which was maintained in subsequent analyses.

In the first model we fit there were no interactions involving iconicity (all $|t| < 1$); most of the two-way interactions were non-significant ($|t| < 1.2$), the only exceptions being Age of learning BSL × Phonological Complexity and Age of learning BSL × Order which were both retained in the next model, along with all main effects. First considering the main effects of control variables varying between items, neither the partial effects of Familiarity nor log(Frequency + 1) reliably predicted naming latencies (Familiarity $\beta = 21$, 95% CI (−48, 90), $t = 0.61$; log Frequency $\beta = 1$, 95% CI (−146, 148), $t = 0.01$). The effect of Order of presentation was reliable (quadratic $\beta = -.0035$, 95% CI (−.0052, −.0018), $t = -4.01$; linear coefficient $|t| < 1$): responses at both the beginning and end of the task were faster than those in the middle. It is not surprising that there was an overall slowdown mid-task given that participants produced signs for 424 pictures (average task length 35 min). Somewhat surprisingly there was no effect of Phonological complexity ($\beta = 28$, 95% CI (−40, 96), $t = .81$), which may suggest that the measures derived from Mann et al. (2010) are inadequate to predict production latencies among fluent signers.[6] There was a reliable effect of Age of learning BSL (β (Late–Early) = 251, 95% CI (77, 425), $t = 2.83$) with RTs for early signers significantly faster compared to late signers. There were two interactions involving Age of learning BSL which were further explored with separate analyses for early and late signers. First, Order of presentation was significant only for late signers (early: quadratic $t = -4.43$; late: quadratic $t = -1.38$), perhaps related to their slower naming latencies overall. Second, although Phonological complexity did not reach significance in either group separately, there was more evidence for its effect in late signers ($\beta = 86$, 95% CI (−12, 185), $t = 1.71$) than for early signers ($\beta = 22$, 95% CI (−46, 90), $t = 0.64$). Crucially, when considering these factors, the partial effect of Iconicity was still a significant predictor of response times ($\beta = -75$, 95% CI (−124, −26), $t = -3.00$). Overall, iconic signs tended to elicit faster responses than less iconic signs (see Figs. 6a and 6b).

We then conducted an additional series of analyses including AoA as an additional control predictor (also including its interaction with Iconicity, and with Age of learning BSL). This model otherwise contained the same terms as the final model above. As with the previous analysis the model failed to converge so we removed the random slopes for Order of Presentation before proceeding. Just as we observed in the main analysis, there were no reliable effects of Frequency, Familiarity or Phonological complexity, nor interactions with Iconicity, and there were reliable main effects of Age of learning BSL (early learners

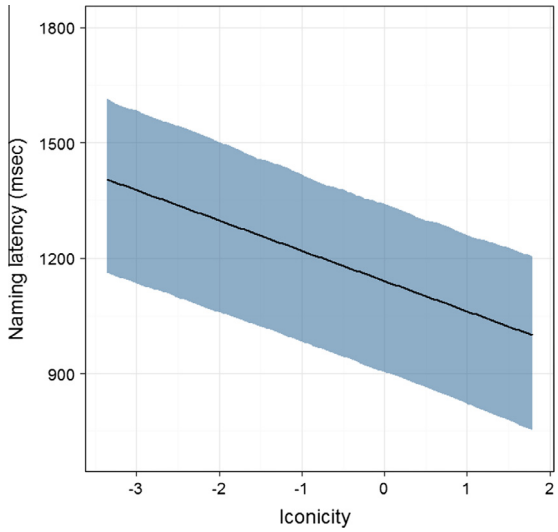

Fig. 6a. (Upper panel): Partial effect of iconicity in the analysis of correct naming latencies in Experiment 3. Solid lines indicate estimate of the mean; shaded areas depict upper and lower quartiles (50% of the predicted naming latencies occur within this range.

faster than late) and Order of presentation (slowest in the middle of the study). The main effect of iconicity was still reliable ($\beta = -79$, 95% CI (−135, −23), $t = -2.78$); qualified by a reliable interaction between iconicity and AoA (interaction $\beta = -37$, 95% CI (−67, −7), $t = -2.45$). There was also a main effect of AoA ($\beta = 48$, 95% CI (7, 90), $t = 2.28$) such that later-learned signs were produced slower. The effect of AoA ratings for signs was not modulated by Age of learning BSL (interaction $\beta = -22$, 95% CI (−54, 11), $t = -1.31$); despite the different learning contexts of early and late signers, AoA ratings predict naming comparably for both.

In order to better understand the nature of the interaction between iconicity and AoA, we carried out follow-up analyses. In the first, we treated AoA as discrete rather than continuous by dividing signs into two groups at the median value for AoA (early-acquired: those rated as learnt before age 7.4, and late-acquired AoA: those rated as learnt after age 7.4; 47 signs in each group). We also removed non-significant predictors (Familiarity and Frequency) from the model above. The interaction between iconicity and AoA was also present when AoA was discretized (interaction $\beta = -92$, 95% CI (−178, −6), $t = -2.09$). We then fit separate models for early- and late-acquired AoA signs, considering only iconicity, signer group and order of presentation as predictors. For early-acquired signs, the effect of iconicity was no longer reliable ($\beta = -32$, 95% CI (−94, 31), $t = -0.99$); but the effect of iconicity was reliable for late-acquired signs ($\beta = -129$, 95% CI (−204, −54), $t = -3.38$) as illustrated in Fig. 7 (upper panels). Signs learnt later in life were affected by iconicity, with more iconic signs being produced faster.[7] We then

[6] We also fit an additional model treating phonological complexity as four separate parameters of handshape, location, movement and hands. This model did not fare any better; none of the individual features' parameter estimates was significant (all $|t| < 1$) and the combined model was not any better than a model with no phonological complexity measure at all.

[7] This difference does not seem to arise as a result of differences in power to detect effects among these item sets, as reliable differences of between early and late signers were observed for both sets of signs and with relatively comparable magnitudes (early-acquired signs β(Late learners–Early learners) = 332, 95% CI (136, 527), $t = 3.327$; late-acquired signs β(Late learners–Early learners) = 244, 95% CI (52, 435), $t = 2.492$).

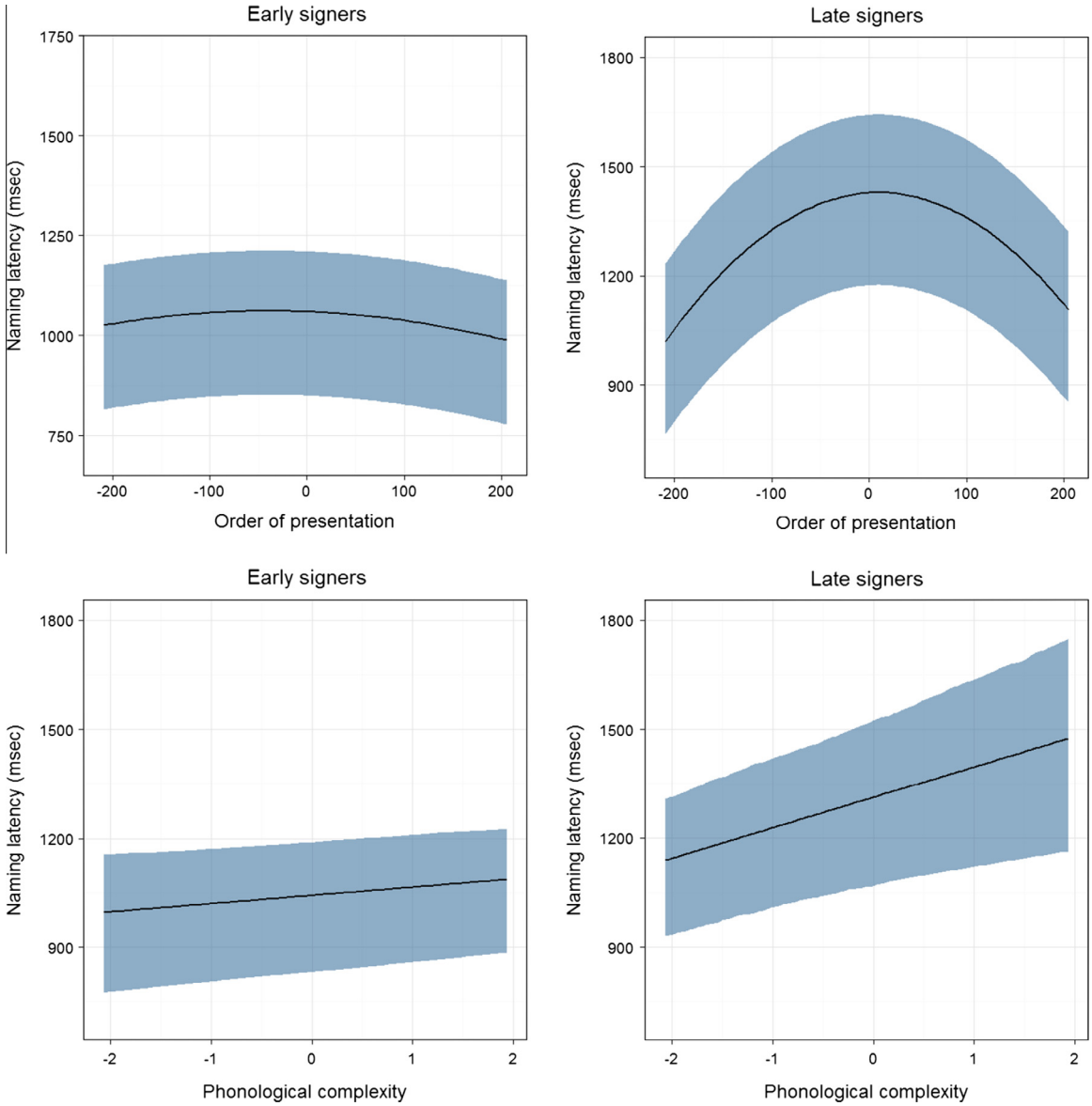

Fig. 6b. Partial effects of significant predictors in the analysis of correct naming latencies in Experiment 3. Solid lines indicate estimate of the mean; shaded areas depict upper and lower quartiles (50% of the predicted naming latencies occur within this range). Upper panel (appearing on previous page): main effect of iconicity. Middle: interaction between age of learning BSL and (centered) order of presentation. Lower: interaction between age of learning BSL and (centered) phonological complexity. Lower and middle panels also reflect the main effect of group, with early signers faster overall

examined this interaction from another angle, instead treating iconicity as a discrete measure (low vs high iconicity, median-split) and evaluating the effects of AoA (continuous measure) for these two groups of signs separately. The effect of AoA was reliable for low-iconicity signs ($\beta = 65$, 95% CI (3, 126), $t = 2.06$): later-acquired signs were named slower than earlier-acquired signs. But for high-iconicity signs, AoA did not affect naming latency ($\beta = 14$, 95% CI (-18, 45), $t = 0.86$); see Fig. 7 (lower panels). Finally, for late-acquired signs only, we fit a model in which signs were distinguished by the type of feature expressed iconically (perceptual, motoric, or both). This factor did not modulate the effect of iconicity we observed for late-learned signs (interaction $|t| < 1.1$).

Discussion

The main finding of Experiment 3 is that iconicity has a sizable facilitatory effect on sign production, but not for all signs: just later-acquired signs showed this effect. For these signs an increase of one unit on the seven-point iconicity rating scale corresponds to nearly 130 ms advantage in naming latency once other factors are taken into account, while there was no reliable effect of iconicity for early acquired signs. In order to understand why this might be, we turn to studies of spoken language where AoA effects on picture naming are well documented, with words learned earlier being named faster even when other

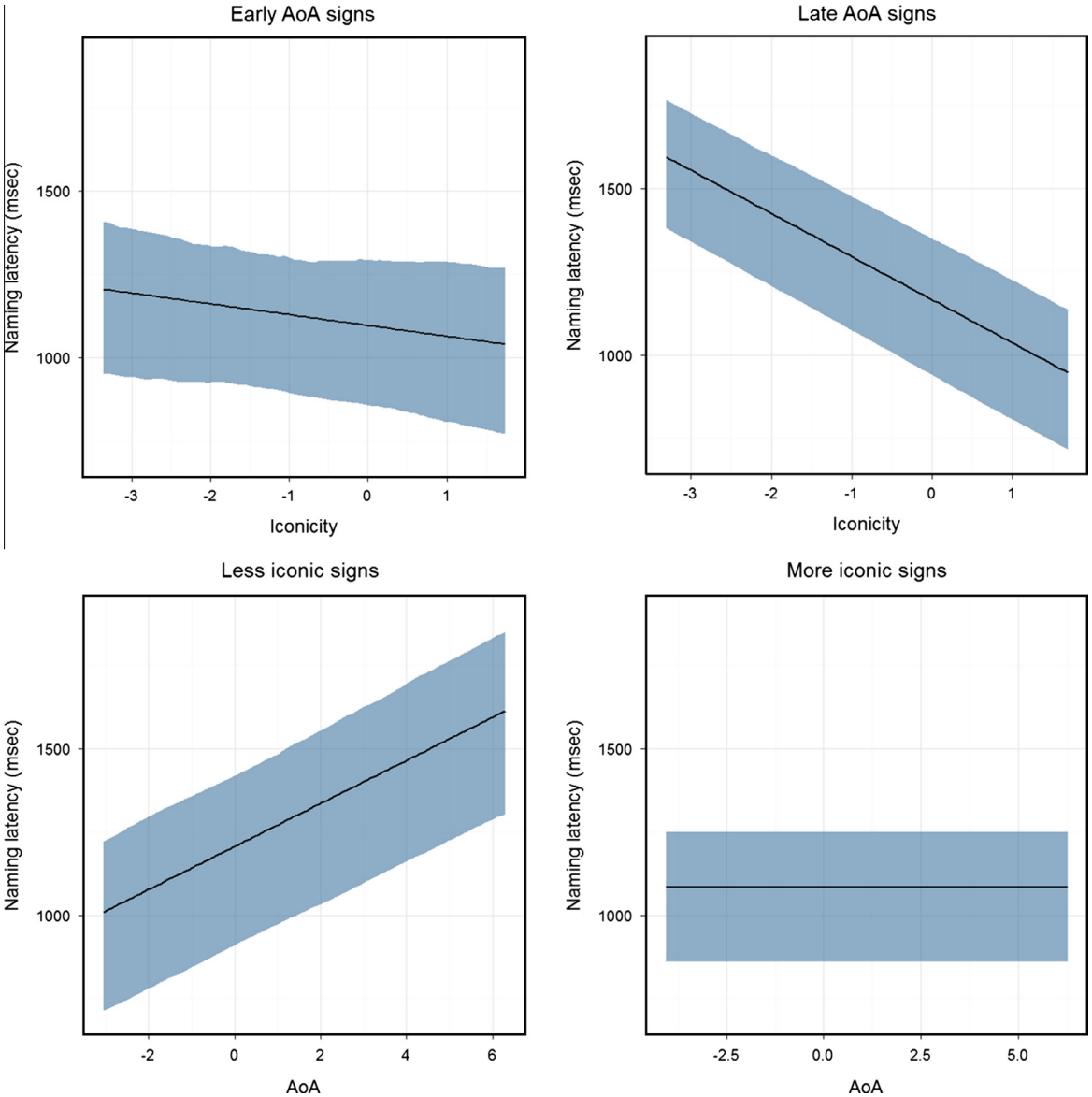

Fig. 7. Main results of the analyses including ratings of sign AoA: Depiction of the interaction between iconicity and AoA in picture naming. Upper panels: partial effect of iconicity (centered) for early-acquired signs (left panel) and late-acquired signs (right panel) on naming latencies. Lower panels: partial effect of AoA for less iconic signs (left panel) and more iconic signs (right panel). Solid lines indicate estimate of the mean; shaded areas depict upper and lower quartiles (50% of the predicted naming latencies occur within this range).

factors are controlled (e.g., Brysbaert & Ghyselinck, 2006; Carroll & White, 1973). Studies exploring this phenomenon in more detail converge on the conclusion that AoA differences affect phonological retrieval (e.g. Belke et al., 2005; Hernandez & Fiebach, 2006; Kittredge et al., 2008; Navarrete et al., 2013); phonological retrieval is slower and more difficult for later-learned words. When we look at the iconicity-AoA interaction with these findings in mind, we see precisely this same pattern for less iconic signs, conforming well to the studies from spoken languages which used highly arbitrary words. But for more iconic signs the disadvantage for later-learned signs was eliminated: the iconic link between conceptual features and phonological form serves to counteract difficulties in retrieving the phonological form. Why was there was no iconicity effect for early acquired signs? There may be different reasons for this. First, it may well be that our study did not have a sufficiently large number of items spanning the range of AoA and iconicity to detect a small but reliable effect of iconicity in early-acquired signs, as was observed developmentally by Thompson et al. (2012) who found that iconicity facilitated acquisition of signs to some extent even in very young children. Alternatively, it may well be that variables such as AoA and iconicity contribute to a

"horse race" toward phonological retrieval and iconicity has greater impact only when retrieval is slower, or when there is some degree of difficulty.

It is also important to note that we did not find effects of frequency or familiarity. After all, if iconicity affects production by facilitating phonological retrieval especially when this is difficult, its effects should also be observed for items that are rare, unfamiliar and phonologically complex as well. However, frequency and familiarity did not vary much in the set. For frequency, hardly any of the items ($N = 29$ out of 92) occurred twice or more in the corpus from which BSL frequency measures were obtained (see Fenlon et al, 2014). For familiarity, the majority of these signs were rated as highly familiar (mean of 5.3 on a 1–7 scale, and highly skewed toward "familiar" such that more than half had familiarity ratings between 5.5 and 6.5); items were also extremely concrete (mean = 4.75 on a 1–5 scale). We can thus consider these variables to be highly controlled among the materials used, minimizing the possibility that effects of iconicity might instead be due to confounds with these factors.

In this study we have found for the first time an effect of iconicity on sign production, allowing us to conclude that iconicity is involved in lexical access for both comprehension and production, rather than coming into play only for specific tasks where it is directly relevant. We also found an effect of group: people who learned to sign early were significantly faster in picture naming than later learners. This finding could be considered as a replication of previous studies of spoken languages in which bilinguals are slower to name pictures in their second language (e.g. Gollan, Montoya, Fennema-Notestine & Morris, 2005; Ivanova & Costa, 2008; Kroll, Bobb & Wodniecka, 2006; Ransdell & Fischler, 1987; see Hanulová, Davidson & Indefrey, 2011 for a review). However, we do not have clear evidence that the late learners of BSL have English as their first language; some of them may have BSL as a first language but with later exposure. Slower responses by such participants would be expected, given other evidence that delayed exposure to a first language may have long-lasting consequences for proficiency even many years later (e.g., Cormier et al., 2012; Ferjan Ramirez, Leonard, et al., 2013; Ferjan Ramirez, Lieberman, et al., 2013; Mayberry & Eichen, 1991; Mayberry & Tuchman, 1985; Meier, 1991). While there was a main effect of group on naming latencies, the effect of iconicity did not vary between groups. Even though it is well attested that late learners report using iconicity as an explicit strategy for language learning (e.g. Baus et al., 2012; Campbell, Martin, & White, 1992) this did not lead to differences in the effects of iconicity on highly-practiced tasks like picture naming.

General discussion

The experiments presented here provide important support to the idea that iconicity affects language processing generally, rather than being a strategy linked to specific tasks. In comprehension, iconic phonological features facilitate the retrieval of corresponding conceptual features. In Experiment 1 we found that BSL signers' responses in picture–sign matching were facilitated when the iconic property of a sign was expressed saliently in the picture, while non-signers showed no benefit in matching such pictures to English words and moreover, we found iconicity effects for both highly typical and less typical semantic features. In Experiment 2, BSL signers' responses to a phonological decision task (indicate direction of movement) were also facilitated by iconicity, showing that facilitative effects can occur on a form-related task where signs' iconic properties are related to the task, further supporting the idea that specific properties of meaning are more highly activated by iconic than noniconic sign forms. Finally, Experiment 3 showed that picture naming is facilitated by iconicity. That the effect of iconicity was limited only to later-acquired signs provides important constraints on the mechanisms by which iconicity affects processing; signs whose forms are easily retrieved do not benefit from iconicity in production. While the possibility remains that some other unknown factor may be responsible for the effects, across the three studies we have ruled out numerous possibilities including: task-related strategies, length, familiarity, frequency, and phonological complexity; we have also found no evidence that iconicity effects are present only for certain kinds of iconic signs and not others.

An interesting puzzle is given by the fact we observed that iconicity effects were modulated by AoA in production (Experiment 3), but not in sign recognition (Experiment 2 and reanalysis of Thompson et al., 2010) AoA effects are well attested in studies of spoken language comprehension (see Brysbaert & Cortese, 2011; Cortese & Khanna, 2007; Hernandez & Li, 2007) with later acquired words showing processing disadvantages, just like in production. These differences may be related to the different demands required by producing vs. comprehending language (e.g. Vigliocco & Hartsuiker, 2002): production, whether spoken or signed, requires retrieval and articulation of a fully phonologically specified form while comprehension requires activation of meaning from any form information. The presence of iconicity effects on production of late-acquired BSL signs, and AoA effects for less iconic signs, suggests that when lexical forms are not as well learnt as early-acquired ones, the system benefits from an additional semantic boost. However, for early-acquired signs (just like for arbitrary words in English), this is not necessary. As a result, while features related to perception and action may activate the phonological features iconically associated with them in production, this activation may only be visible in cases of difficulty that slow down the process. Alternatively, one could argue that iconic links are present only for less well-established lexical forms (late-acquired, less-frequent) but are no longer necessary once phonological form is more solidly established. However, this latter possibility is unlikely, because we found no difference between native and late learners of BSL in the effects of iconicity on picture naming. If iconic links are only in place during learning, we would expect native signers to show reduced effects even for later-learned signs. Thus, we consider iconicity as a property that increases the strength of the connection between

semantic and phonological features bidirectionally; with the presence of iconicity effects in the production of late-acquired signs indicating that it may be especially important when retrieval is more demanding and therefore the stronger link between meaning and form can provide a useful additional boost. Finally, we should be cautious about the extent to which our results apply to iconicity beyond concrete domains (e.g. Taub, 2001). While we have ruled out explanations of iconicity effects based entirely on concreteness, by selecting highly concrete items and taking concreteness ratings into account where available, the effects we observed may not speak to metaphorical iconicity, or other aspects of iconicity for abstract concepts.

Our results have implications for both spoken and signed languages, and there is some evidence supporting the notion that in spoken languages too, iconic relationships between meaning and form affect online processing (Ohtake & Haryu, 2013; Westbury, 2005; Meteyard et al., 2015). Given that current embodied cognition theories are calling for more explicit, testable theories (e.g. Fischer & Zwaan, 2008; Glenberg et al., 2013; Meteyard et al., 2012) iconicity effects whether in spoken or signed language offer new constraints concerning how to model the relationship between meaning and form in both comprehension and production. One possibility is that models need to be altered to incorporate modality-specific representations, or dual pathways for iconic forms, rather than amodal representations that mediate between meaning and form.

However, it may be possible to accommodate iconicity effects within some standard models of processing, if we consider iconicity just as a form of regularity. That is, traditional (non-embodied) models that include featural representations of meaning, with intermediate abstract representations that map onto wordform features, can establish stronger links for regular mappings between meaning and form via changes in connection weights on the basis of experience across multiple exemplars. The particular type of iconicity we examined in Experiment 2 (upward/downward phonological movement in BSL corresponding to physical upward/downward movement in the sign's meaning) might be considered a highly regular mapping. This is because movement along a vertical axis is highly likely to be represented by vertical phonological movement, even though many exceptions can be identified (particularly signs with vertical phonological movement that does not correspond to directional movement in the referent). Under such a view, the effects of iconicity may be no different from aspects of regularity that are not iconic, such as the association between the English past tense and "-ed" (see Plaut et al., 1996). An account of iconicity based solely on regularity, however, does not appear to accommodate effects of iconic properties that depict distinctive properties rather than shared mappings across many exemplars, as was the case for many of the items in Experiment 1, such as antlers for DEER in BSL, peeling motion for BANANA, or the flickering of a flame for CANDLE (see Appendix B for more examples). Moreover, models of this type depend on learning for the strengthening of links that eventually facilitate processing. As such, a regularity-based account also appears inconsistent with developmental evidence suggesting a central role of iconic mappings in bootstrapping language learning (e.g. Asano et al., 2015; Imai & Kita, 2014; Laing, 2014), hinging upon mappings between real-world experience and linguistic forms that are not just statistical associations.

Why would languages be iconic at all?

Although theories of language representation and processing embed a core assumption of arbitrariness, Perniss et al. (2010) argue that this is essentially a historical accident, a consequence of the development of language processing theories around western European languages which happen to be particularly impoverished in iconic forms (at least at the lexical level). When a broader range of languages is considered, many spoken languages are revealed to have far more iconic words than English or other western European languages, iconicity that reflects experience well beyond that of an acoustic nature, encompassing not only other dimensions of perceptual experience but also motoric, affective and temporal properties of entities and events (see Perniss et al., 2010 for a review). Once language is considered in communicative contexts, the presence of iconicity in spoken languages is far greater: reflected, for example, in prosody and co-speech gesture (typically ignored in most studies of language processing) which offer substantial additional means of iconic expression. So, even though iconicity is more immediately evident as a property of signed languages, especially in contrast to "impoverished" languages like English, it is perhaps unwarranted to discard iconicity as a quirk of modality.

If iconicity is widely present across languages and plays a role in processing, how then can this be reconciled with the obvious characteristics of arbitrariness present across all languages whether spoken or signed? Perniss et al. (2010) discuss this in terms of the very different roles played by arbitrariness and iconicity acting in tandem to fulfill different language functions. Arbitrariness has been argued to play an important role in communicative effectiveness; arbitrary forms permit finer distinctions to be made among relatively similar entities (Gasser, 2004; Gasser, Sethuraman & Hockema, 2005; Monaghan, Christiansen & Fitneva, 2011), distinctions that may be difficult or impossible to make with a purely iconic vocabulary in conditions of information loss (e.g. speech in noisy environments). Iconicity is instead argued to play a foundational role in providing a link between linguistic forms and physical experience with the world, both in terms of processing and language learning (see also Thompson, 2011; Thompson et al., 2012 for discussion of the timing of iconicity effects in development), in line with studies demonstrating interplay between language and other aspects of perception and cognition (e.g. embodied cognition; see Barsalou, 1999, 2009; Meteyard, Rodriguez Cuadrado, Bahrami & Vigliocco, 2012). Gasser et al. (2005) argue further that the particular communicative demands related to different semantic domains may also drive the relative contribution of iconicity and arbitrariness: a greater level of arbitrariness should be present in

categories where fine-grained meaning distinctions are crucial (e.g. nouns referring to artifacts or food items) and greater level of iconicity should be present where such fine-grained distinctions are not so essential (e.g. expressives related to motion and other aspects of physical activity).

How might effects of iconicity come to be?

If indeed iconicity serves a foundational role in linking linguistic forms to meaning, iconic mappings should not just affect adult processing but also learning. Despite the still-standard claim that iconic signs do not seem to be overly represented in children's earliest signs (Morgan et al., 2006; Orlansky & Bonvillian, 1984), and children's errors in signing do not exhibit overly iconic properties (Meier, 1982, 1991; Meier et al., 2008) recent evidence suggests that iconicity actually does facilitate acquisition even among very young children. Specifically, a recent language acquisition study found a relationship between iconicity and the signs comprehended and produced by very young children learning to sign at home, even as young as 11–20 months (Thompson, Vinson, Woll & Vigliocco, 2012). Thompson et al. analysed data from the BSL version of the MacArthur Bates Communication Development Inventory (Woolfe et al., 2010), which consisted of parental reports of their child's sign comprehension and production (ages 11–30 months). For BSL signs for which iconicity ratings are available (Vinson et al., 2008), signs with higher iconicity were more likely to be comprehended or produced even by very young children. This effect of iconicity increased for older children (greater magnitude for children aged 21–30 months than for children aged 11–20 months). We also know that by the time child signers become adult signers they have developed full awareness of iconicity, making use of it in areas such as poetry and humor (see Kaneko & Sutton-Spence, 2012; Sutton-Spence & Napoli, 2012).

While the results described by Thompson et al. are correlational in nature, there also exist studies for which causation can be claimed. In one such study, Imai et al. (2008) created novel verbs which were judged by adults (Japanese- and English-speaking) to be sound-symbolic of particular actions, and other novel verbs that were not sound-symbolic. They then used these novel verbs in a learning task with three-year-old children. The children showed an advantage in learning the novel sound-symbolic words, compared to the novel words that were not sound-symbolic, suggesting that iconic mappings of this nature can facilitate early language development, particularly where actions are concerned. Such early effects of iconicity suggest that some aspects of iconic mappings between concepts and phonology facilitate learning even before children can explicitly or obviously appreciate iconicity. Why then are iconic signs not more highly represented in children's early language? One important consideration is that not all aspects of sign phonology are iconic of the referent, and not all visible properties of a given referent are reflected in sign phonology; most iconic signs include some arbitrary elements. Taking this argument one step further, Perniss and Vigliocco (2014) propose that in language development, iconicity plays the critical role of reducing referential ambiguity such that children can "tune in" to the correct referents of linguistic labels (see also Imai & Kita, 2014).

Conclusions

In the present work we have shown that iconicity has general effects on comprehension and production, as a direct consequence of interplay between meaning and iconic forms: in particular the link between features related to perception and action, and phonological features. While this effect sometimes manifests as an interfering effect of iconicity (as in some phonological decisions and translation tasks) it can still be explained in terms of this general characteristic of iconic forms. Such findings are difficult to accommodate within standard language processing frameworks which were developed on the basis of languages relatively impoverished in iconic lexical forms (e.g. English and other western European languages) and therefore embed arbitrariness as a core principle of meaning-form mappings. The iconicity effects we observe point toward a more central role of iconicity in language processing than has been previously considered: linking language to bodily experience. Crucially we argue that these findings should not only be limited to sign language processing but should equally apply to iconic forms in spoken languages as well.

Acknowledgments

This work was supported by the Economic and Social Research Council of Great Britain (Grants RES-620-28-6001 and 620-28-6002: Deafness, Cognition and Language Research Centre, and Grant ES/K001337/1 to D. Vinson), by the European Union (FP6-2004-NESTPATH, Grant 028714 to G. Vigliocco) and by a Birmingham University Fellowship awarded to R. Thompson. Trial-level data from the experiments reported here can be downloaded from the UK Data Service: doi: 0.5255/UKDA-SN-851078.

Appendix A. Details of item selection, Experiment 1

A.1 Selection of experimental items

Because we wanted to select items for which the features expressed iconically in BSL varied in typicality, we started with all of the picturable items included in the English semantic feature norms of McRae et al. (2005). From this set of pictures we first excluded those where BSL signers tended to name the picture with (a) a fingerspelled form (in which handshapes representing English letters are used to spell out the word) rather than a lexical sign, (b) a non-iconic sign, (c) a more generic term than the corresponding word in English (or vice versa), or (d) an alternative iconic sign that highlights different properties of the object (all based on the judgments of R.S., a native BSL signer). Because we needed two different pictures for each sign, one with the iconic property expressed saliently,

and one in which the iconic property was not salient, another set of items were also excluded where it was not possible to produce both types of pictures (e.g. BALL, for which the BSL sign reflects its round shape). For all of the remaining signs, we selected a number of different line drawings, taken when possible from Thompson et al. (2009), and others from sources like Snodgrass and Vanderwart (1980), Druks and Masterson (2000), Szekely et al. (2004) and Dell'Acqua, Lotto and Job (2000), as well as a number of commercial clip art packages. Additional pictures were drawn to order by the same artist who produced pictures for Thompson et al. (2009). We divided candidate signs into high and low typicality categories by consulting the English semantic feature norms of McRae et al. (2005). Signs were considered to be of high typicality if any semantic feature corresponding to the iconic property in BSL was produced by one-third or more of McRae's subjects – signifying that particular semantic feature is highly salient to English speakers. Otherwise a sign was considered to be of low typicality.

We then embarked on a multi-stage norming procedure, gathering name agreement norms from both BSL signers and non-signing English speakers. Subjects (12 BSL signers in the initial phase, and 14 English speakers) were presented with each picture and asked to name it (some of the first BSL signers we tested were also asked whether they could think of alternative signs that might be regularly used by our potential subjects). In addition to assessing name agreement, we paid close attention to the iconic features produced by BSL signers to ensure that they produced our intended iconic feature. In a number of cases, follow-up questions (especially to BSL signers) indicated that name agreement problems were occurring not due to the language itself, but due to particulars of specific pictures (e.g., a picture of a mouse with a prominent nose (the iconic feature in BSL) which was mistaken for other rodents). We therefore created additional versions of some pictures, and alterations of others, and carried out a second round of name agreement norming in both BSL (n = 12) and English (n = 10).

A.2. Selection of filler items

In addition to the 54 experimental signs/words and their corresponding pictures, we also selected a large number of filler items (pictures and signs/words). Because the crucial experimental signs occurred twice in the experiment, each time paired with a different picture, we included many filler trials which sampled from one of two different pictures corresponding to one sign or word (but crucially not varying in the salience of an iconic property). Some of these were matching trials (where the picture matched the sign or word), others were mismatching trials, so that the appearance of a second picture referring to a word or sign (no picture appeared more than once in the experiment) could not serve as a cue to the correct response before the sign or word actually appeared. Other filler pictures and signs/words occurred only once in the experiment, including many that were presented for the first time only in the second block of the experiment, in order to further disguise the purpose of the experiment as well as balancing the number of match and mismatch responses.

Appendix B

Details of the items reported in analyses of Experiment 1. ProdFreq1: Production frequency of iconic feature according to norms from McRae et al. (2005). ProdFreq2: Production frequency of second listing for iconic feature (if present). Distinct: Feature distinctiveness measure, 1/(number of concepts for which that feature was listed in McRae et al., 2005). A value of 1 indicates that no other word in the set shared that feature. For features not listed in McRae et al.'s norms, distinctiveness is marked "NA".

Concept	Iconic feature	Typicality	ProdFreq1	ProdFreq2	Distinct
aeroplane	wings/flight	High	25	20	0.022
ambulance	flashing lights	High	18		1
axe	chopping	High	29	23	0.5
banana	peeling	High	10	7	0.25
belt	waist/thickness/wearing	High	9	8	0.5
bike	pedaling	Low	9		0.25
bread	slicing	Low	8		1
cake	shape	Low	0		NA
candle	flickering flame	High	15	9	1
car	steering	High	12		1
cat	whiskers	High	11		0.125
caterpillar	shape/movement	High	11		0.1
chair	armrests	Low	10		0.333
clock	hand/clock face/movement	High	18	7	1
cow	horns	Low	0		NA
crown	shape/worn on head	High	20		1

(*continued on next page*)

Appendix B (*continued*)

Concept	Iconic feature	Typicality	ProdFreq1	ProdFreq2	Distinct
deer	antlers	High	13		0.25
eagle	beak	High	12		0.026
elephant	trunk	High	23		1
fox	nose	Low	0		NA
frog	bulging neck	Low	0		NA
goat	beard	High	14		1
gorilla	beating chest	Low	0		NA
gun	shape	Low	0		NA
helicopter	propeller/motion	High	16	11	0.5
house	roof	Low	9		0.143
kite	holding string	High	21		0.333
lion	paws	Low	0		NA
motorbike	twisting handlebars	Low	0		NA
mouse	nose	Low	0		NA
necklace	neck/thickness/shape	High	23		0.167
owl	large eyes	High	16		0.063
peg	clipping	Low	10		0.333
piano	playing	High	27		0.143
pig	snout	High	12		1
pipe (smoking)	holding	Low	0		NA
plug	plugs into wall	High	8	6	1
rabbit	ears	High	14		0.2
ring	placing on finger	High	20		0.5
saxophone	hold/press keys	Low	7		0.143
seal	flippers	High	12		1
skateboard	flat surface/movement	Low	8		1
spider	legs/movement	High	19	5	1
telephone	handset	Low	10		1
tent	shape/ropes	Low	0		NA
toaster	pops up	Low	9		1
trolley	holding/pushing	Low	0		NA
umbrella	opening movement	Low	0		NA

Appendix C. Norming signs for motion iconicity (Experiment 2)

In order to test specifically whether iconic motion in BSL signs affected decisions it was necessary to obtain a measure directly reflecting whether the motion in a sign reflected iconic properties (in contrast to the more general measure of iconicity which provides an index of the extent to which *any* parameter of a sign reflects properties of the referent). To do this we closely followed the methods of Meteyard and Vigliocco (2009) who collected similar norms for English verbs. Each BSL sign was presented in video format, along with a visual display depicting four possible directions of motion, and a fifth category for "no directional motion" (see Fig. A1). For each sign, participants were asked to judge whether the *meaning* of the sign refers to motion in any of the particular directions depicted: downward, upward, to the right, to the left, or no motion.

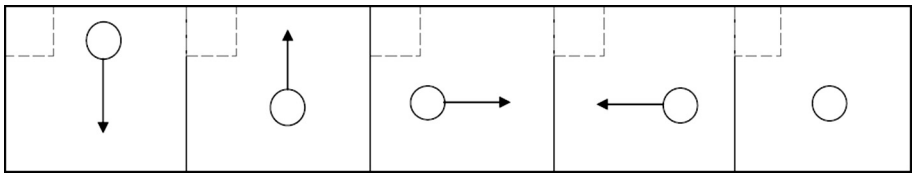

Fig. A1. Visual display accompanying BSL signs for motion iconicity norming. Adapted from Meteyard and Vigliocco (2009).

Table A1
Signs used in Experiment 2 and their motion iconicity ratings, along with the average scores for iconic motion in each direction.

Gloss motion	Phonological motion	Motion iconicity	Down	Up	Left	Right	Null
AEROPLANE	UP	3.9	0.7	3.9	1.1	0.1	0
AFTERNOON	DOWN	0	0	0.4	0	0	3.9
APPLE	UP	1.5	0.4	1.5	0.3	0	2.4
ARRIVE	DOWN	1.1	1.1	0.8	1.8	0.4	0.8
ART	DOWN	1.8	1.8	0.6	0.8	0.5	2
AUDIENCE	UP	1.5	0.4	1.5	0.7	0.2	2
BALL	DOWN	1	1	1.5	1.4	0.5	2
BANANA	UP	0.6	0.8	0.6	0.4	0.3	2.4
BLACK	DOWN	0	0	0.4	0	0	3.9
BOOT	UP	2	0.7	2	0	0	1.8
BORN	DOWN	3.5	3.5	0.3	0	0	0.8
BOTTLE	UP	0.8	0	0.8	0	0	3.5
BOWL	UP	0.7	0	0.7	0.4	0.2	3.2
BUY	DOWN	0.9	0.9	0.8	0.2	0	3.2
CAN	UP	2.7	0	2.7	0.3	0	1.9
CHAMPAGNE	UP	3.5	0.3	3.5	0.8	0.3	1.1
CHEF	UP	1.2	0	1.2	0	0	3.1
CLIMB	UP	3.9	0.7	3.9	0.3	0.2	0
CLOCK	DOWN	1.6	1.6	2.3	2.7	0.4	0.4
CONNED	DOWN	0.8	0.8	0.2	0	0	3.4
COPY	DOWN	0.3	0.3	0.8	0.4	0.3	2.8
CORKSCREW	UP	2.8	0.4	2.8	0.3	1.1	0
CRY	DOWN	3.7	3.7	1.5	0	0	0.2
CULTURE	UP	0.7	0.7	0.7	0.2	0.1	3.2
DECIDE	DOWN	1	1	0	0	0	3.6
DEER	UP	1.8	0.4	1.8	1.6	0.6	1.9
DIVE	DOWN	3.1	3.1	0.7	0.5	0.8	0
DONT_BELIEVE	UP	0.4	0.3	0.4	0	0.4	3.5
DONT_LIKE	UP	0.8	0.4	0.8	0.6	0.2	3.1
DOWN	DOWN	3.9	3.9	0.4	0	0	0
DRESS	DOWN	2.2	2.2	0.4	0	0	2.3
DRINK	UP	3.1	1.1	3.1	0.3	0	0.4
DROP	DOWN	3.5	3.5	0.4	0	0	0.7
DROWN	DOWN	3.8	3.8	0.4	0	0	0.4
EXTRA	UP	0.8	0	0.8	1.5	0.3	2
FAKE	DOWN	0.3	0.3	0	0	0	4
FEATHER	UP	0.8	0.2	0.8	0.4	0.6	2.8
FINALLY	DOWN	0.9	0.9	0.4	0	0	3.6
FIREWORKS	UP	3.9	0.5	3.9	0.9	0.8	0.4
FIZZ	UP	2.7	0.7	2.7	0	0	1.2
FOOTBALL	UP	3.5	1.3	3.5	0.7	0.5	1
GIRAFFE	UP	2.1	0.3	2.1	0.6	0.4	2.6
GO_PAST_HEAD	UP	1.2	0.4	1.2	0	0	2.4
GOD	UP	2.3	0	2.3	0	0	2
HALF	DOWN	1.6	1.6	0	0.3	0.4	2
HAT	DOWN	2.2	2.2	0.4	0	0	2.3
HELICOPTER	UP	3.6	0.1	3.6	0.6	0.2	0.4
IDEA	UP	1.8	0	1.8	0	0	3.1
INJECTION	DOWN	1.2	1.2	0.4	1.5	0	1.5
JACKET	DOWN	1.8	1.8	0	0.7	0.5	2.3
JUMPER	DOWN	1.7	1.7	0.4	0	0	2.7
LETTER	DOWN	1.1	1.1	0.4	0.3	0	2.8
LIFT	UP	4	0.6	4	0	0	0
LIFT_RIDE	UP	2.7	1	2.7	0.6	0.1	0.8
LUCKY	DOWN	0.4	0.4	0	0.4	0.3	3.5
MAN	DOWN	0.4	0.4	0.4	0	0	3.5
MEASURE	UP	2.3	0.9	2.3	0.6	0.5	1.2
MOUNTAIN	UP	2.4	0	2.4	0.6	0.4	1.6
NEVER	DOWN	1.1	1.1	0.4	0	0	2.8
NOW-I-GET-IT	DOWN	0.4	0.4	0	0	0	3.6
NUT	UP	0.4	0	0.4	0	0	3.9
OLD	DOWN	0.6	0.6	0.8	0	0	3.5
PARIS	UP	2	0	2	0.3	0.2	2
PARK	UP	0.4	0	0.4	0.4	0.3	3.5
PLEASE	DOWN	0.7	0.7	0.4	0	0	3.8
POLICE	UP	0.4	0	0.4	0.4	0.6	3.2

(*continued on next page*)

Table A1 (*continued*)

Gloss motion	Phonological motion	Motion iconicity	Down	Up	Left	Right	Null
POOR	UP	0.4	0.4	0.4	0	0.3	3.4
POUND	DOWN	0	0	0.4	0	0	3.9
PROFESSIONAL	UP	0	0	0	0	0	3.6
PSYCHOLOGY	DOWN	0.8	0.8	0	0.3	0	3.2
PUB	DOWN	0.6	0.6	0.8	0	0	2.8
RAIN	DOWN	3.7	3.7	1.8	0	0	0
RELY	DOWN	1	1	0.4	0	0	3.2
RESEARCH	DOWN	0.6	0.6	0.7	0	0	3.2
RIGHT	DOWN	0.4	0.4	0	0	0	3.6
ROCKET	UP	3.5	0	3.5	0.3	0	0.8
RUDE	UP	0.4	0	0.4	0	0	3.9
SENSITIVE	UP	0.8	0.4	0.8	0	0.3	2.8
SHAVE	DOWN	2.3	2.3	1.2	0.6	0	1.1
SHOWER	DOWN	3.8	3.8	0.8	0	0	0
SKILL	UP	0.7	0	0.7	0.4	0.4	2.8
SLEEP	DOWN	1.5	1.5	0.4	0.4	0	2
SLOW	UP	1.4	0.7	1.4	0.3	0.4	2.3
SLY	DOWN	0.7	0.7	0.4	0	0	3.5
SMILE	UP	2.9	0.4	2.9	1.3	0.9	1.1
SOUP	UP	2.4	0.6	2.4	0	0.3	1.2
SPORT	DOWN	1.2	1.2	1.4	0.4	0.4	2.4
STAMP	DOWN	2.1	2.1	0.8	0.2	0	1.9
STRONG	UP	1.1	0	1.1	0	0	3.2
SUBMARINE	DOWN	3.5	3.5	0.7	2	0.1	0.4
SUNRISE	UP	3.6	0	3.6	0.3	1.3	0.8
SWALLOW	DOWN	3.8	3.8	0.4	0	0	0.4
TAKE	UP	1.5	1.5	1.5	0	0	1.6
TEA	UP	1.9	0.4	1.9	0.3	0	2.5
TENT	DOWN	1.4	1.4	0.8	0.3	0.2	2.4
THUNDER	DOWN	3.1	3.1	1.2	1.2	0.8	0.4
TIRED	DOWN	0.7	0.7	0	0	0	3.9
UMBRELLA	UP	3.6	0	3.6	0	0	0
UP	UP	4	0	4	0	0	0
WALES	DOWN	0	0	0.4	0	0.4	3.2
WALL	UP	1.9	0.6	1.9	0	0	2.4
WANT	DOWN	0.6	0.6	0.8	0	0	3.5
WATERFALL	DOWN	3.1	3.1	1.2	1.3	0.3	0
WHALE	UP	3.3	0.6	3.3	0.2	0.5	1.5
WORLD	DOWN	0.7	0.7	1.5	1.2	0.7	2

Examples were provided to make it clear that ratings should reflect *iconic* directional motion rather than vertical and/or horizontal motion which is only phonological in nature. Participants were asked to provide numeric rankings to reflect relative importance/salience of the different directions, in the event that they felt more than one direction was appropriate (for example, BOUNCE reflects both downward and upward motion). A ranking of 1 indicated that a particular direction was most important for that sign, 2 next most important, and so on. If a direction of motion was considered to be not relevant at all, it should be left blank. Rankings were transformed by subtracting them from a value of 5 and then averaged across participants for each item, so that directional scores for each of the five possibilities ranged from a maximum of 4 (all participants rated the same direction as maximally important for a given sign) to a minimum of 0 (no participants rated that direction as important for that sign).

Our measure of motion iconicity was determined by the score assigned to the direction of a sign's phonological movement, made possible by our selection of signs which had clear movement either upwards or downwards. As a result this measure reflects motion iconicity specifically on the vertical dimension. For example, the sign SWALLOW (downward phonological movement) had a downward score of 3.8, upward score of 0.4, leftward score of 0, rightward score of 0 and null-movement score of 0.4. Therefore 3.8 was assigned as the motion iconicity score for SWALLOW. In a similar way, the sign CHAMPAGNE (upward phonological movement) had a downward score of 0.3, upward score of 3.5, leftward score of 0.8, rightward score of 0.3 and null-movement score of 1.1 – thus receiving a motion iconicity score of 3.5. Average motion iconicity values for each sign used in the experiment appear in Table A1. We used these values in the regression models described in the main text.

Appendix D. Items used in Experiment 3 and their characteristics

Gloss	Familiarity	Iconicity	AoA	Log (frequency + 1)
AEROPLANE	6.1	6.4	4.6	1.15
APPLE	5.8	5.9	na	0.90
BANANA	4.7	5.2	na	0.00
BELT	5.6	6.7	7.2	0.00
BIKE	5.9	5.7	5.2	0.60
BINOCULARS	4.4	6.3	8.2	0.00
BOMB	5.3	4.9	11.3	0.30
BOOT	5.8	5.1	8.6	0.00
BOTTLE	6.0	5.4	6.7	0.48
BOWL	5.5	6.3	na	0.30
BOX	6.1	6.0	6.4	0.78
BOY	6.4	1.9	4.9	1.11
BREAD	6.5	4.0	5.8	0.30
CANOE	4.5	5.8	10.9	0.00
CARDS	5.6	5.9	7.8	0.60
CASTLE	4.9	4.9	8.3	0.60
CHAIR	6.0	3.9	6.9	0.70
CHEESE	6.0	2.5	8.1	0.00
CHEF	4.6	5.0	na	0.00
CHERRY	5.1	3.0	10.3	0.00
CHOCOLATE	6.0	1.9	7.3	0.00
CHURCH	5.7	4.3	8.2	1.00
CLOCK	4.9	6.2	6.8	0.30
CLOTHESPIN	3.6	4.9	10.6	0.00
CLOUD	4.9	5.4	6.2	0.00
COOKER	6.1	5.6	5.8	0.00
CORKSCREW	3.6	4.7	14.0	0.00
CROCODILE	5.2	6.2	5.6	0.00
DEER	4.8	6.0	8.3	0.00
DOG	6.1	3.6	4.6	1.34
DRESS	5.2	4.7	5.0	0.70
DRILL	4.1	5.2	10.1	0.00
DUCK	5.3	6.0	4.5	0.00
EARTH	6.4	6.4	na	0.00
EGG	4.6	3.4	6.9	0.00
ENGLAND	6.5	2.0	9.2	1.20
FEATHER	3.3	4.8	na	0.00
FIRE	5.3	5.9	6.2	1.18
FLOWER	5.4	3.5	5.7	0.00
FOOTBALL	5.7	4.9	8.1	1.30
GIRAFFE	2.9	5.6	na	0.00
GIRL	4.1	3.4	9.9	0.00
GOLD	5.6	2.9	8.5	0.00
HAIRBRUSH	5.6	6.5	4.7	0.30
HAMMER	4.7	5.6	6.7	0.00
HEARING-AID	5.7	5.9	10.2	0.00
HELICOPTER	4.7	5.6	11.3	0.00
HOUSE	5.8	4.5	5.9	1.66
ICE-CREAM	5.9	6.0	3.6	0.78
IRON	5.2	6.3	6.8	0.00
JUMPER	6.0	5.5	5.6	0.00
KANGAROO	4.2	5.5	8.9	0.00
KEY	6.3	6.1	5.9	0.30
KNIFE	5.9	5.2	7.8	0.00

(*continued on next page*)

Appendix D. Items used in Experiment 3 and their characteristics (*continued*)

Gloss	Familiarity	Iconicity	AoA	Log (frequency + 1)
LAUGH	6.3	5.1	7.2	1.36
LETTER	5.9	4.2	7.9	0.30
LOCK	5.9	5.4	7.8	0.70
MAN	6.9	3.1	na	1.40
MONKEY	5.5	5.9	4.7	0.00
MOON	5.5	6.3	7.6	0.00
MOUNTAIN	4.5	5.8	na	0.00
MOUSE	5.2	2.3	6.7	0.30
NUT	5.5	3.4	9.4	0.30
PAINTBRUSH	5.2	5.3	5.6	0.00
PAPER	5.9	1.6	6.4	0.00
PILLOW	4.3	4.8	8.0	0.00
POOL	5.3	6.3	9.4	0.48
POOR	5.5	2.7	8.9	0.00
POTATO	1.7	1.3	11.5	0.00
RABBIT	5.5	6.0	4.5	0.00
RAKE	2.6	4.4	12.0	0.00
RAZOR	5.6	6.7	na	0.00
SANDWICH	6.5	4.9	8.5	0.00
SAW	4.3	5.0	8.1	0.00
SCARF	4.4	5.0	9.2	0.30
SCHOOL	4.9	1.6	9.1	1.68
SHOP	6.2	1.9	8.5	0.90
SHOWER	7.0	6.2	na	0.00
SKI	5.5	6.6	8.4	0.00
SKIRT	5.4	6.5	6.3	0.00
SMILE	6.3	6.1	na	0.00
STRAWBERRY	3.1	2.1	9.6	0.85
SWING	4.7	6.0	4.9	0.00
TENT	3.8	5.9	na	0.60
TIE	5.5	6.4	7.3	0.00
TREE	5.7	6.3	6.3	0.95
TURTLE	2.9	5.0	13.9	0.00
UMBRELLA	5.8	5.7	na	0.00
VIOLIN	4.8	5.7	9.9	0.00
WALL	5.6	6.3	na	0.60
WATER	6.5	2.0	7.3	0.90
WRITE	6.3	6.5	5.3	1.40

References

Aristotle (350BCE/1941). Categoriae [Categories] (E.M. Edghill, Trans.). In R. McKeon (Ed.), *The basic works of Aristotle* (pp. 7–37). New York: Random House.

Armstrong, D. F., & Wilcox, S. E. (2009). Gesture and the nature of semantic phonology. *Sign Language Studies, 9*, 410–416.

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex, 63*, 196–205.

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. *Memory & Cognition, 40*, 83–92.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–609.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London: Biological Sciences, 364*, 1281–1289.

Bates, D., Bolker, B., Maechler, M., & Walker, S. (2013). *Linear mixed-effects models using Eigen and S4*. <http://cran.r-project.org/web/packages/lme4 Version 1.0-4> Downloaded 08.09.13.

Baus, C., Carreiras, M., & Emmorey, K. (2012). When does iconicity in sign language matter? *Language and Cognitive Processes*. http://dx.doi.org/10.1080/01690965.2011.620374.

Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition, 96*, B45–B54.

Bosworth, R., & Emmorey, K. (2010). Effects of iconicity and semantic relatedness on lexical access in American Sign Language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1573–1581.

Bremner, A., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). Bouba and kiki in Namibia? Western shape-symbolism does not extend to taste in a remote population. *Cognition, 126*, 165–172.

Brown, W. (1915). Practice in associating color names with colors. *Psychological Review, 22*, 45–55.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109*, 204–223.

Brown, R., & McNeill, D. (1966). The "tip-of-the-tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*, 325–337.

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology, 64*, 545–559.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency-related, partly frequency-independent. *Visual Cognition, 13*, 992–1011.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*, 904–911.

Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108–135). Cambridge, MA: MIT Press.

Campbell, R., Martin, P., & White, T. (1992). Forced choice recognition of sign in novice British Sign Language learners. *Applied Linguistics, 13*, 185–201.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture naming latency. *Quarterly Journal of Experimental Psychology, 25*, 85–95.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11*, 63–65.

Collins, A. C., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Cormier, K., Schembri, A., Vinson, D., & Orfanidou, E. (2012). First language acquisition differs from second language acquisition in prelingually deaf signers: Evidence from grammatical processing of British Sign Language. *Cognition, 124*, 50–65.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical decision performance above and beyond 22 other predictor variables: An analysis of 2340 words. *Quarterly Journal of Experimental Psychology, 60*, 1072–1082.

Davis, R. (1961). The fitness of names to drawings: A cross-cultural study in Tanganyika. *British Journal of Psychology, 52*, 259–268.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*, 801–838.

Dell'Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers, 32*, 588–615.

Dove, G. (2013). Intermediate representations exclude embodiment. *Behavioral and Brain Sciences, 36*, 353–354.

Druks, J., & Masterson, J. (2000). *An object and action naming battery*. London: Psychology Press.

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Erlbaum and Associates.

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society B, 369*, 20130301. http://dx.doi.org/10.1098/rstb.2013.0301.

Emmorey, K., & Corina, D. (1990). Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills, 71*, 1227–1252.

Fenlon, J., Schembri, A., Rentelis, R., Vinson, D., & Cormier, K. (2014). Using conversational data to determine lexical frequency in British Sign Language: The influence of text type. *Lingua, 143*, 187–202. http://dx.doi.org/10.1016/j.lingua.2014.02.003.

Ferjan Ramirez, N., Leonard, M. K., Torres, C., Hatrak, M., Halgren, E., & Mayberry, R. I. (2013). Neural language processing in adolescent first-language learners. *Cerebral Cortex, 2013*. http://dx.doi.org/10.1093/cercor/bht137.

Ferjan Ramirez, N., Lieberman, A. M., & Mayberry, R. I. (2013). The initial stages of language acquisition begun in adolescence: When late looks early. *Journal of Child Language, 40*, 391–414.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language – A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology, 61*, 825–850.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.

Fromkin, V. (Ed.). (1973). *Speech errors as linguistic evidence*. The Hague: Mouton.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*, 455.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.). *Language production: Speech and talk* (Vol. 1). London: Academic Press.

Garrett, M. F. (1984). The organization of processing structure for language production: Application to aphasic speech. In D. Caplan, A. R. Lecours, & A. Smith (Eds.), *Biological perspectives on language* (pp. 172–193). Cambridge, MA: MIT Press.

Garrett, M. F. (1992). Lexical retrieval processes: Semantic field effects. In E. Kittay & A. Lehrer (Eds.), *Frames, fields and contrasts: New essays in semantic and lexical organization*. Hillsdale, NJ: Erlbaum.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 231–256). Amsterdam: North Holland.

Gasser, M. (2004). The origins of arbitrariness in language. In *Proceedings of the annual conference of the cognitive science society*, Chicago, IL.

Gasser, M., Sethuraman, N., & Hockema, S. (2005). Iconicity in expressives: An empirical investigation. In S. Rice & J. Newman (Eds.), *Conceptual structure, discourse, and language 7: Empirical and experimental methods in cognitive/functional research* (pp. 1–18). Stanford, CA: CSLI Publications.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Wiley Interdisciplinary Review of Cognitive Science, 7*, 155–170.

Glaser, W. R. (1992). Picture naming. *Cognition, 42*, 61–105.

Glenberg, A. M., & Kaschak, M. P. (2003). The body's contribution to language. *Psychology of Learning and Motivation, 43*, 93–126.

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the revolution to embodiment: 25 years of cognitive psychology. *Perspectives on Psychological Science, 8*, 573–585.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*, 1220–1234.

Grosjean, F. (1981). Sign and word recognition: A first comparison. *Sign Language Studies, 32*, 195–219.

Grote, K., & Linz, E. (2003). The influence of sign language iconicity on semantic conceptualization. In W. G. Mueller & O. Fischer (Eds.), *From sign to signing: Iconicity in language and literature 3* (pp. 23–40). Amsterdam: Benjamin.

Halle, M. (1991). Phonological features. In W. Bright (Ed.), *International Encyclopaedia of Linguistics*. Oxford, UK: Oxford University Press.

Hamano, S. (1998). *The sound-symbolic system of Japanese*. Stanford, California: CSLI Publications.

Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes, 26*, 902–934.

Hawthorne, J. W. (1934). An attempt to measure certain phases of speech. *The Journal of General Psychology, 10*, 399–414.

Hernandez, A. E., & Fiebach, C. J. (2006). The brain bases of reading late learned words: Evidence from functional MRI. *Visual Cognition, 13*, 1027–1043.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*, 638.

Hohenberger, A., Happ, D., & Leuninger, H. (2002). Modality-dependent aspects of sign language production. Evidence from slips of the hands and their repairs in German Sign Language (Deutsche Gebärdensprache (DGS)). In K. Cormier, R. Meier, & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken language* (pp. 112–142). Cambridge, Mass: Cambridge University Press.

Hohenberger, A., & Waleschkowski, E. (2005). Language production errors as evidence for language production processes. The Frankfurt Corpora. In S. Kepser & M. Reis (Eds.), *Linguistic evidence. Empirical theoretical and computational perspectives* (pp. 285–305). Mouton de Gruyter.

Holland, M. K., & Wertheimer, M. (1964). Some physiognomic aspects of naming, or, maluma and takete revisited. *Perceptual and Motor Skills, 19*, 111–117.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Proceedings of the Royal Society B: Biological Sciences, 369*, 20130298. http://dx.doi.org/10.1098/rstb.2013.0298.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition, 109*, 54–65.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127*, 277–288.

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007). What do English speakers know about gera-gera and yota-yota?: A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education around the Globe, 17*, 53–78.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 130.

Johnson, R. C. (1967). Magnitude symbolism of English words. *Journal of Verbal Learning and Verbal Behavior, 6*, 508–511.

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin, 120*, 113–139.

Kakehi, H., Tamori, I., & Schourup, L. (1996). *Dictionary of iconic expressions in Japanese*. Berlin: Walter de Gruyter.
Kaneko, M., & Sutton-Spence, R. (2012). Iconicity and metaphor in sign language poetry. *Metaphor and Symbol, 27*, 107–130.
Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science, 35*, 575–586.
Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics, 35*, 379–415.
Kita, S. (2001). Semantic schism and interpretive integration in Japanese sentences with a mimetic: A reply to Tsujimura. *Linguistics, 39*, 419–436.
Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture naming errors. *Cognitive Neuropsychology, 25*, 463–492.
Klima, E., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.
Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.
Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140*, 14.
Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition, 114*, 19–28.
Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition, 9*, 119–135.
Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 92.
Kunihira, S. (1971). Effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior, 10*, 427–429.
Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. *First Language, 34*, 387–405.
Levelt, W. J. L. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–38.
MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.
Mann, W., Marshall, C. R., Mason, K., & Morgan, G. (2010). The acquisition of sign language: The impact of phonetic complexity on phonology. *Language Learning and Development, 6*, 60–86.
Mayberry, R. I., & Eichen, E. B. (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language, 30*, 486–512.
Mayberry, R., & Tuchman, S. (1985). Memory for sentences in American Sign Language: The influence of age of first sign learning. In W. Stokoe & V. Volterra (Eds.), *Proceedings of the IIIe. International symposium on sign language research* (pp. 120–125). Silver Spring, Md: Linstock Press.
McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavioral Research Methods, Instruments, and Computers, 37*, 547–559.
Meier, R. P. (1982). *Icons, analogues, and morphemes: The acquisition of verb agreement in American Sign Language*. Unpublished doctoral dissertation, University of California, San Diego.
Meier, R. P. (1991). Language acquisition by deaf children. *American Scientist, 79*, 60–70.
Meier, R. P., Mauk, C. E., Cheek, A., & Moreland, C. J. (2008). The form of children's early signs: Iconic or motoric determinants? *Language Learning and Development, 4*, 63–98.
Meteyard, L., Rodriguez Cuadrado, S., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex, 48*, 788–804.
Meteyard, L., & Vigliocco, G. (2009). Verbs in space: Axis and direction of motion norms for 299 English verbs. *Behavior Research Methods, 41*, 565–574.
Meteyard, L., & Vigliocco, G. (2008). The role of sensory and motor information in semantic representation. In P. Calvo & A. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 291–312). London: Elsevier Academic Press.
Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. *Neuropsychologia*. http://dx.doi.org/10.1016/j.neuropsychologia.2015.01.042.
Monaghan, P., & Christiansen, M. H., (2006). Why form-meaning mappings are not entirely arbitrary in language. In *Proceedings of the 28th annual conference of the cognitive science society*, Vancouver.
Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General, 140*, 325–347.
Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1152.
Morgan, G., Barrière, I., & Woll, B. (2006). The influence of typology and modality on the acquisition of verb agreement morphology in British Sign Language. *First Language, 26*, 19–43.
Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development, 5*, 37–57.
Navarrete, E., Scaltritti, M., Mulatti, C., & Peressotti, F. (2013). Age-of-acquisition effects in delayed picture-naming tasks. *Psychonomic Bulletin & Review, 20*, 148–153.
Newcombe, P. I., Campbell, C., Siakaluk, P. D., & Pexman, P. M. (2012). Effects of emotional and sensorimotor knowledge in semantic processing of concrete and abstract nouns. *Frontiers in Human Neuroscience, 6*. http://dx.doi.org/10.3389/fnhum.2012.00275.
Newkirk, D., Klima, E. S., Canady Pedersen, C., & Bellugi, U. (1980). Linguistic evidence from slips of the hand. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand* (pp. 165–197). New York: Academic Press.
Newman, S. S. (1933). Further experiments in phonetic symbolism. *The American Journal of Psychology, 45*, 53–75.
Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 65*, 115.
Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115*, 357–395.
Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*, 299–325.
Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition, 112*, 181–186.
Oda, H. (2000). *An embodied semantic mechanism for mimetic words in Japanese*. Ph.D. dissertation, Indiana University.
Ohtake, Y., & Haryu, E. (2013). Investigation of the process underpinning vowel-size correspondence. *Japanese Psychological Research, 55*, 390–399.
Oldfield, R. C., & Wingfield, A. (1964). The time it takes to name an object. *Nature, 202*, 1031–1032.
Orlansky, M., & Bonvillian, J. D. (1984). The role of iconicity in early sign language acquisition. *Journal of Speech and Hearing Disorders, 49*, 287–292.
Ormel, E., Gijsel, M. A. R., Hermans, D., Bosman, A. M. T., Knoors, H., & Verhoeven, L. (2010). Semantic categorization: A comparison between deaf and hearing children. *Journal of Communication Disorders, 43*, 347–360.
Ormel, E., Hermans, D., Knoors, H., & Verhoeven, L. (2009). The role of sign phonology and iconicity during sign processing: The case of deaf children. *Journal of Deaf Studies and Deaf Education, 14*, 436–448.
Ortega, G., & Morgan, G. (2010). Comparing child and adult development of a visual phonological system. *Language, Interaction and Acquisition, 1*, 67–81.
Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.
Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
Paivio, A. (1991). *Images in mind: The evolution of a theory*. New York: Harvester Wheatsheaf.
Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. New York: Lawrence Erlbaum.
Paivio, A., Yuille, J. C., & Smythe, P. C. (1966). Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology/Revue canadienne de psychologie, 20*, 362.
Parault, S. J., & Parkinson, M. (2008). Sound symbolic word learning in the middle grades. *Contemporary Educational Psychology, 33*, 647–671.
Parault, S. J., & Schwanenflugel, P. J. (2006). Sound-symbolism: A piece in the puzzle of word learning. *Journal of Psycholinguistic Research, 35*, 329–351.
Perniss, P. M. (2012). Use of sign space. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 412–431). Berlin: Mouton de Gruyter.
Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages.

Frontiers in Psychology, 1, 227. http://dx.doi.org/10.3389/fpsyg.2010.00227.
Perniss, P. M., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B, 369*. http://dx.doi.org/10.1098/rstb.2013.0300.
Pitcher, B. J., Mesoudi, A., & McElligott, A. G. (2013). Sex-biased sound symbolism in English-language first names. *PLoS One, 8*, e64825. http://dx.doi.org/10.1371/journal.pone.0064825.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review, 103*, 56.
Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22*, 253–279.
R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies, 8*, 3–34.
Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language, 26*, 405.
Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*, 460–499.
Roca, I. M. (2003). Phonology. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science* (pp. 636–645). London: Nature Publishing Group.
Roelofs, A. (1993). A spreading-activation theory of lemma retrieval in speaking. In W. J. M. Levelt (Ed.), *Lexical access in speech production* (pp. 107–142).
Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). The structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review, 111*, 205–235.
Rosch, E. H. (1973). Natural categories. *Cognitive Psychology, 4*, 328–350.
Rosch, E., & Mervis, C. B. (1975). Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605.
Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.
Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge, UK: Cambridge University Press.
Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology, 12*, 225.
Schembri, A. (2003). Rethinking "classifiers" in signed languages. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 3–34). Mahwah, New Jersey: Lawrence Erlbaum Associates.
Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, 114–126.
Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6*, 174–215.
Stokoe, W. C. (1991). Semantic phonology. *Sign Language Studies, 71*, 107–114.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.
Sutton-Spence, R., & Napoli, D. J. (2012). Deaf jokes and sign language humor. *Humor, 25*, 311–337.
Sutton-Spence, R., & Woll, B. (1999). *The linguistics of British Sign Language: an introduction*. Cambridge UK: Cambridge University Press.
Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andanova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*, 247–250.
Tarte, R. D., & Barritt, L. S. (1971). Phonetic symbolism in adult native speakers of English: Three studies. *Language and Speech, 14*, 158–168.
Taub, S. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.
Taylor, I. K. (1963). Phonetic symbolism re-examined. *Psychological Bulletin, 60*, 200.
Thompson, R. L. (2011). Iconicity in language processing and acquisition: What signed languages reveal. *Language and Linguistics Compass, 5*, 603–616.
Thompson, R. L., Emmorey, K., & Gollan, T. H. (2005). Tip-of-the-fingers experiences by ASL signers: Insights into the organization of a sign-based lexicon. *Psychological Science, 16*, 856–860.
Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Quarterly Journal of Experimental Psychology, 64*, 2392–2404.
Thompson, R. L., Vinson, D. P., Fox, N., & Vigliocco, G. (2013). Seeing the world through a visual language: Visual world in British Sign Language. *Proceedings of the Cognitive Science Society*.
Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2009). The link between form and meaning in American Sign Language: Lexical processing effects. *Journal of Experimental Psychology: Language, Memory, and Cognition, 35*, 550–557.
Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British Sign Language: Lexical processing effects in a phonological decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 550–557.
Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science, 23*, 1443–1448.
Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science, 8*, 314–317.
Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin, 128*, 442–472.
Vigliocco, G., & Kita, S. (2006). Language-specific effects of meaning, sound and syntax: Implications for models of lexical retrieval in production. *Language and Cognitive Processes, 21*, 790–816.
Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition, 1*, 219–247.
Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). The meanings of object and action words. *Cognitive Psychology, 48*, 422–488.
Vigliocco, G., Vinson, D. P., Martin, R. C., & Garrett, M. F. (1999). Is "count" and "mass" information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language, 40*, 534–558.
Vigliocco, G., Vinson, D. P., Woolfe, T., Dye, M. W., & Woll, B. (2005). Words, signs and imagery: When the language makes the difference. *Proceedings of the Royal Society B, 272*, 1859–1863.
Vinson, D. P. (2009). *Representing meaning: A feature-based model of object and action words*. Unpublished PhD thesis, University College London. <http://discovery.ucl.ac.uk/14891/>.
Vinson, D. P., Cormier, K., Denmark, T., Schembri, A., & Vigliocco, G. (2008). The British Sign Language (BSL) norms for age of acquisition, familiarity and iconicity. *Behavior Research Methods, 2008*, 1079–1087.
Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods, 40*, 183–190.
Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1256.
Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language, 93*, 10–19.
Woolfe, T., Herman, R., Roy, P., & Woll, B. (2010). Early vocabulary development in deaf native signers: A British Sign Language adaptation of the communicative development inventories. *Journal of Child Psychology and Psychiatry, 51*, 322–331.