Iconicity as a general property of language: evidence from spoken and signed languages

Pamela Perniss, Robin L. Thompson1 and Gabriella Vigliocco*

Deafness, Cognition and Language Research Centre, Cognitive, Perceptual and Brain Sciences Research Department, Division of Psychology and Language Sciences, University College London, London, UK

Edited by:
Morten H. Christiansen, Cornell University, USA

Reviewed by:
Benjamin Bergen, University of California San Diego, USA
Sherman Wilcox, University of New Mexico, USA

*Correspondence:
e-mail: g.vigliocco@ucl.ac.uk
1Pamela Perniss and Robin L. Thompson have contributed equally to the paper. Order of names determined by coin toss.

Current views about language are dominated by the idea of arbitrary connections between linguistic form and meaning. However, if we look beyond the more familiar Indo-European languages and also include both spoken and signed language modalities, we find that motivated, iconic form-meaning mappings are, in fact, pervasive in language. In this paper, we review the different types of iconic mappings that characterize languages in both modalities, including the predominantly visually iconic mappings found in signed languages. Having shown that iconic mapping is present across all languages, we then proceed to review evidence showing that language users (signers and speakers) exploit iconicity in language processing and language acquisition. While not discounting the presence and importance of arbitrariness in language, we put forward the idea that iconicity need also be recognized as a general property of language, which may serve the function of reducing the gap between linguistic form and conceptual representation to allow the language system to “hook up” to motor, perceptual, and affective experience.

Keywords: iconicity, sound-symbolism, spoken and signed languages

ARBITRARINESS IN LANGUAGE: THE RECEIVED VIEW

The relation between words and real-world referents has intrigued scholars since antiquity. Early debates centered on the origin of words (as names for things), specifically, on the nature of their relation to the things they stand for. In Plato’s Cratylus, the oldest documented of these debates, Socrates is asked to contemplate the question of whether names belong to their objects “naturally” or “conventionally.” The latter of these two possibilities, namely that form and meaning are linked by convention and tradition alone, has come to dominate our modern thinking about language. Words, and more generally language as a symbolic system, are conceived of as being arbitrarily related to the world. At the lexical level, the phonological form of a linguistic sign is considered to have no relationship to its meaning.

The idea of an arbitrary connection between form and meaning is generally associated with Saussure, but, in fact, John Locke established a firm foothold for the idea much earlier when he argued in his Essay Concerning Human Understanding (1690) that the existence of different languages (with very different words for the same objects) is evidence against a “natural” connection between form and meaning. Since everyone perceives the world in the same way, there should be only one human language if properties of objects could in any way (i.e., by “natural” connection) determine the names given them. The persuasive force of this argument, and its far-reaching impact, is not difficult to fathom.

Today, no one would subscribe to the idea of an actual “natural” connection between linguistic signs and their denotata. The idea that each object could have an inherently “correct” name known from the object itself strikes us as antiquated and arcane. Instead, the current mainstream view has embraced the polar opposite idea: namely that convention alone determines the relationship between form and meaning. Indeed, if we look at the lexicon of English (or that of other Indo-European languages), we might be forgiven for thinking that there could be anything but a conventionally determined, arbitrary connection between a given word and its referent. For the vast majority of English words there is an arbitrary relationship between form and meaning. There is nothing, for example, in the sequence of sounds represented by /haus/ that indicates that it refers to “a building for human habitation” (house) in English.

Language is understood to be a system of conventional symbols shared by a community of users. Users of English must agree about the meaning of /haus/, and this knowledge must be passed on from generation to generation. We can see the stamp of conventionalization even in the case of onomatopoeia, i.e., words like buzz or meow, which imitate sounds in the real world (usually made by objects or animals). The English expression for the sound a rooster makes (cock-a-doodle-doo) is quite different from the German expression (kikiriki), which is different again from the French expression (cocorico), and all of which are arguably quite different from the actual sound which emerges from a rooster.

The evidence for conventionalized, arbitrary form-meaning mappings as defining our system of communication seems almost self-evident. Moreover, it is consistent with, and surely engendering of, the idea that language is a wholly symbolic system, the elements of which are manipulated on an abstract level of representation. It is this idea that defines and dominates current, received views about the acquisition, production, and comprehension of language, as, for example, in the strict modular separation of levels for lemma (semantic/grammatical representation) and lexeme (phonological representation) in some models of lexical retrieval and production (e.g., Garrett, 1984; Levelt, 1989; Levelt et al., 1999). According to these authors, not only are meaning and form represented and
retrieved separately in production, but they are fully independent. Whereas the last 20 years have seen numerous debates concerning the degree of interaction and overlap between semantic and syntactic processing in models of word and sentence comprehension, the question of whether and to what extent phonological and semantic representations overlap has received almost no attention. One need only read, for example, the opening paragraphs of the highly influential Behavioral and Brain Sciences article by Levelt et al. (1999) on speech production to realize just how fundamental the assumption of arbitrariness is in theories of language processing. For Levelt et al. (1999), language acquisition and use is set into motion through the merging of two separate systems: “[W]ord production emerges from a coupling of two initially independent systems, a conceptual system and an articulatory motor system. This duality is never lost in the further maturation of our word production system.” (p. 1)

This “major rift” in the system, as Levelt et al. (1999) call it, renders impossible the presence of anything like a “natural” connection, the more contemporary version of which we may frame in terms of iconicity or motivation, between form and meaning. In this paper, we argue that the preoccupation with arbitrariness in language has eclipsed a proper acknowledgment of non-arbitrary, i.e., iconic and motivated form-meaning mappings in language. When we widen the scope of linguistic inquiry to include larger expanses of the globe, beyond Indo-European languages (and thus more diverse linguistic structures), as well as language in both modalities (spoken and signed), it is clear that many languages in fact make widespread and systematic use of non-arbitrary form-meaning mappings.

In Section “Iconicity in Languages: Is It Really There?,” we show that regular correspondences between form and meaning exist across both spoken and signed language modalities, motivated by perceptuo-motor properties of real-world experiences. We use iconicity as a blanket term for a broad range of phenomena, including what has been referred to in the literature as sound-symbolism, mimetics, ideophones, and iconicity. It is our view that these different phenomena (discussed below), while varying in amount and degree, all have in common some non-arbitrary, iconic mapping between form and meaning. In providing a comprehensive overview of the presence of iconicity in language, we challenge the received view that language is fundamentally arbitrary in nature: a view which is dismissive of iconic form-meaning mappings as being negligible in quantity and linguistically inferior in quality. After demonstrating the presence and availability of iconic mappings in language, we go on in Section “Are we Sensitive to Iconicity?” to review research findings that indicate that iconicity has a role to play in processing. The findings show that language users are in fact sensitive to iconicity, suggesting that iconic, motivated mappings influence language processing in a way that calls into question the deep “rift” between levels of form and meaning. We further discuss the possible role that iconicity may play in lexical acquisition. Finally, in Section “General Discussion,” we discuss the implications of considering iconicity, and not only arbitrariness, as a ubiquitous, fundamental feature of language.

**ICONICITY IN LANGUAGES: IS IT REALLY THERE?**

Far from constituting a ‘vanishingly small’ proportion of words in any given language (Newmeyer, 1992), iconic mappings can be observed both at the lexical and structural level across modalities, and, importantly, these mappings can be considered to be a general property of languages. For example, well documented is the iconicity of onomatopoeia (the group of words referred to by Newmeyer). Moreover, iconicity in the structure of sentences (reflected in morphosyntax and especially word order) is also well acknowledged in the literature (Haiman, 1980, 1985; Givón, 1985; Berlin, 1994; Croft, 2003). Here, as Croft (2003, p. 102) puts it, we find that “the structure of language reflects in some way the structure of experience.” Examples of such iconicity include that of sequence, contiguity, repetition, quantity, complexity, and cohesion (see also Greenberg, 1963; Haiman, 1980; Croft, 1990; Givón, 1991; Newmeyer, 1992; Levinson, 2000). For example, the principle of “iconicity of sequence” (or “sequential order”) holds that the sequence of forms conforms to the sequence of experience, as in the famous collocation *veni, vidi, vici*. The principle of “iconicity of contiguity” (or “linguistic proximity”) assumes that forms that belong together semantically will occur closer together morphosyntactically than forms that are semantically unrelated (cf. Bybee, 1985). While this type of iconicity has been investigated primarily for spoken languages, and hardly at all for signed languages, we may assume that similar kinds of structural iconicity influence syntagmatic relations in both language modalities.

At the lexical level, however, iconicity in spoken languages has hardly been acknowledged beyond onomatopoeic words, in line with the received view that arbitrariness governs the lexical interface between meaning and form in language (the rift mentioned above). The situation is different for signed languages, where the existence of iconic links between form and meaning cannot escape the eye. Nonetheless, guided by the assumptions of the received view, the role of iconicity has been downplayed. Below we review the literature for spoken and signed languages separately.

**ICONICITY IN SPOKEN LANGUAGES**

In spoken languages, depictive forms that make use of iconic mappings were described as early as the seventeenth century, and noted in more earnest in grammars (e.g., of African languages) dating to the middle of the nineteenth century (Dingemanse, p.c.). However, their seemingly idiosyncratic forms and behavior, as well as their depictive and expressive properties defied standard linguistic description. Within current theoretical perspectives, iconicity is mostly ignored, based mainly on the argument that iconicity in spoken language is only observed for a small number of words. Indeed, direct imitative form-meaning mappings in spoken languages are only possible for acoustic properties and events, which are arguably far less frequent than visual properties and events in our experience. Thus, onomatopoeic words make up only a small portion of spoken languages’ total lexica, although they appear consistently across all spoken languages.

The sounds imitated in onomatopoeia are typically animal sounds (*meow, moo, oink*) or the sounds made by objects in motion or upon impact on other objects (*whoosh, swish, whack, crack, crash, bang*). Some of these words (particularly those referring to object motion and impact) also exhibit another type of iconic mapping called *phonesthesis*. In phonesthemes, a similarity of form, typically in word-initial or word-final consonant clusters, is correlated with a similarity of meaning. For example, in English, words ending in *-ack*, as in *whack* and *crack* above, denote forceful, punctuated
contact, while words beginning with gl- (as in gleam, glow, glint, glitter) denote a meaning related to light of low intensity, and words beginning with wr- (as in writhe, wriggle, wrist, write) refer to twisting (see Firth, 1930 for many more examples). As the overlap between some of these onomatopoeic and phonesthetic forms indicates, the presence of phonesthesia in language may in part be the result of generalization from acoustically iconic form-meaning mappings.

The “small inventory” argument for iconic forms in spoken languages may hold for most Indo-European languages, and has likely been expanded to language on the whole as a consequence of a Euro-centric linguistic perspective. When we move outside the Indo-European language family, however, we find that iconic mappings are prevalent and are used to express sensory experiences of all kinds. Languages for which a large iconic, or sound-symbolic, lexicon has been reported include virtually all sub-Saharan African languages (Childs, 1994), some of the Australian Aboriginal languages (Alpher, 2001; McGregor, 2001; Schulze-Berndt, 2001), Japanese, Korean, Southeast Asian languages (Difflloth, 1972; Watson, 2001), indigenous languages of South America (Nuckolls, 1996), and Balto-Finnic languages (Mikone, 2001).

In addition to more direct acoustic links, these iconic, sound-symbolic mappings evoke sensory, motor, or affective experiences or characterize aspects of the spatio-temporal unfolding of an event. In fact, the majority of sound-symbolic words refer to events or states in which sound is not essential. That is, properties of experiences – including visual, tactile, as well as mental and emotional experiences – may systematically correspond to properties of vowels and consonants, and their patterns of combination (e.g., reduplication) (Hamano, 1998). To give the reader some idea of the rich system of iconically motivated words found in many languages, we give examples in Table 1 of iconic ideophones from Siwu (a Kwa (Niger-Congo) language spoken in eastern Ghana) (Dingemanse, in press) and in Table 1 of iconic mimetics from Japanese (taken from Vigliocco and Kita, 2006).

The examples from both languages make clear the wide range of sensory events that can be incorporated into sound-symbolic mappings: visual sensations, tactile sensations, different types of motion events, manners of motion, physiological states, and psychological states. The examples also make clear that there are systematic form-meaning correspondences in these words that make use of what has been referred to as Gestalt iconicity (Bühler, 1934; see also Dingemanse, in press); the structure of the word can be structurally iconic of the spatio-temporal structure of an event. For example, in both Siwu and Japanese, the full reduplication of syllables map onto events that are iterated or distributed. In Siwu, in particular, unitary events tend to be expressed with monosyllabic forms, and the addition of a lengthened vowel evokes a unitary, but durative event. In Japanese, in particular, the voicing of an initial consonant indicates the size of the object involved (as in the difference between goro and koro in Table 1). Such form-meaning relationships are not always fully productive, but are shared among many mimetic words.

The inventories of sound-symbolic lexical forms in languages which have them are typically substantial. One dictionary of Japanese mimetic words lists more than 1,700 entries (Atoda and Hoshino, 1995). As should be evident from the examples above, mimetic words cannot be considered to be an isolated phenomenon in these languages (as one may argue regarding onomatopoeia in languages such as English or Italian). Sound-symbolic, iconic words play an integral part in the language as adverbials, predicative nominals, verbs, and adjectives. They are used frequently in everyday conversation, and are especially frequent in narratives and story-telling, to help bring to life events through vivid depiction and enactment. In Japanese, they are used in many different forms of established verbal arts (Schourup, 1993), from comic books to novels by Nobel-Prize winning authors.

The inventories of such iconic forms in individual languages are, of course, conventionalized within each language, yet there exist what seem to be universal tendencies in the types of mappings that occur across languages. For example, as discussed further in Section “Are we Sensitive to Iconicity?,” back vowels and voiced consonants

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Examples of sound-symbolism in Siwu and Japanese.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SIWU IDEOPHONIC WORD</strong></td>
<td><strong>Meaning</strong></td>
</tr>
<tr>
<td>Kgbebebe</td>
<td>Rigid posture of a muscular person</td>
</tr>
<tr>
<td>Kpa’kpa’kpa</td>
<td>The jerky walk of a turtle</td>
</tr>
<tr>
<td>Gidgi</td>
<td>Running energetically</td>
</tr>
<tr>
<td>Nyemere-nyemere</td>
<td>Slithering movement</td>
</tr>
<tr>
<td>Tsokwe-tsokwe</td>
<td>Sawing movement</td>
</tr>
<tr>
<td>Bintiri</td>
<td>Sowed at close intervals</td>
</tr>
<tr>
<td>Munyemunjye</td>
<td>Sparkling light</td>
</tr>
<tr>
<td>Saia</td>
<td>Cool sensation</td>
</tr>
<tr>
<td>Yii</td>
<td>The roaring feeling of vertigo</td>
</tr>
<tr>
<td>Yal’al’al’al’</td>
<td>Flowing quietly without obstruction</td>
</tr>
<tr>
<td>Kala’ala</td>
<td>Looking attentively</td>
</tr>
<tr>
<td>Dza</td>
<td>Sudden appearance</td>
</tr>
<tr>
<td>Wil’i</td>
<td>Bright flash</td>
</tr>
<tr>
<td>Wurufu’u</td>
<td>Fluffy texture</td>
</tr>
<tr>
<td>Tsarara</td>
<td>The fluidity of a liquid substance (e.g., palm oil)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>JAPANESE MIMETIC WORD</strong></th>
<th><strong>Meaning</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Goro</td>
<td>Heavy object rolling</td>
</tr>
<tr>
<td>Koro</td>
<td>Light object rolling</td>
</tr>
<tr>
<td>Pota</td>
<td>Thin/little liquid hitting</td>
</tr>
<tr>
<td>Boto</td>
<td>Thick/much liquid hitting</td>
</tr>
<tr>
<td>Gorogoro</td>
<td>A heavy object rolling repeatedly</td>
</tr>
<tr>
<td>Korokoro</td>
<td>A light object rolling repeatedly</td>
</tr>
<tr>
<td>Nununuru</td>
<td>The tactile sensation caused by slimy substance</td>
</tr>
<tr>
<td>Pika</td>
<td>A flash of light</td>
</tr>
<tr>
<td>Pyon</td>
<td>A swift jump</td>
</tr>
<tr>
<td>Yochiyochi</td>
<td>The manner of walking</td>
</tr>
<tr>
<td>Tobotobo</td>
<td>A sluggish manner of walking</td>
</tr>
<tr>
<td>Tekutefu</td>
<td>A brisk manner of walking</td>
</tr>
<tr>
<td>Kutakuta</td>
<td>A state of fatigue</td>
</tr>
<tr>
<td>Sowasowa</td>
<td>The restless anxiety before an important event</td>
</tr>
</tbody>
</table>
tend to evoke large, heavy, and more rounded things, while front
vowels and voiceless consonants evoke smaller, more jagged things
(Köhler, 1929; Ramachandran and Hubbard, 2001). This univer-
sality may be due to the origin of sound-symbolism in imitative
connections between the shape of the mouth (producing sounds)
and the shape of real-world entities (the denotata of the sounds),
as well as possibly a reflection of (imitative) kinesic muscle activity
triggered by certain events/actions (Foroni and Semin, 2009).

ICONICITY IN SIGNED LANGUAGE

Signed languages, produced and perceived in the visual-spatial
modality, obey the same grammatical constraints and linguistic
principles found in spoken languages (for reviews, see Emmorey,
2002; Sandler and Lillo-Martin, 2006). The visual nature of the
modality, however, results in an abundance of direct iconic, visual-
to-visual mappings. Much of what we communicate about is visu-
ally perceived (e.g., where things are, where they are going, how
they are interacting, and what they look like), and the visual-spatial
modality affords a visually iconic depiction of such information
through the placement of the hands (as primary articulators) in
the space in front of the body (i.e., the signing space).

At the phonological level any one sign is made up of features
from three main parameters: the shape of the hands (handshape),
the location of the hands in space or on the body (location),
and the movement of the hands (movement). Iconicity may,
but need not be, present in one or more of these parameters.
While many signs exhibit iconicity, as the modality exploits the
potential for visual iconic mappings, there are also signs which
we may consider to be fully arbitrary, where a visual motivation
for the form-meaning mapping is not apparent. Most signs, in
fact, exhibit both iconic and arbitrary features, and the degree
of iconicity or arbitrariness ascribed to individual signs is best
understood as a continuum. Examples of iconic and non-iconic
signs in British Sign Language (BSL) are given in Figures 1A–D
below. The iconic sign for cry (in Figure 1A) is made with two
extended index fingers (handshape), which move in an alternating
pattern downward from the eyes (movement) on the signer’s face
(location). The sign visually depicts the path of tears as they fall.
In the BSL sign meaning aeroplane (in Figure 1B), the extended
thumb and pinky visually depict an aeroplane’s wings, and the
movement of the hand high in space a straight trajectory depicts
the path motion of flying. Conversely, there is no iconic mapping
between form and meaning for the BSL signs for battery and
afternoon (Figures 1C,D).

Similarly to the Gestalt iconic correspondences discussed
for spoken language above, where the nature of an event can
be reflected in the morphophonological structure of the word,
sign forms can also be modified to reflect the spatio-temporal
structure of events. For example, in the sign forms, like in the
spoken word forms, reduplicated movement patterns indicate
iteration or continuation, with the type of reduplication (e.g.,
short, punctuated vs. smooth, continuous repetitions) provid-
ing additional information. Such modifications can be used to
indicate aspectual information, or mark plurals (see Figure 2;
Klima and Bellugi, 1979). Their iconic, motivated character is
identified more clearly in Wilbur’s (2003, 2009) elaboration of the
Event Visibility Hypothesis. Specifically, Wilbur argues that the

\[ \text{FIGURE 1 | Examples of iconic signs meaning cry (A) and aeroplane (B) and non-iconic signs meaning battery (C) and afternoon (D) in BSL.} \]
Thus, whereas in spoken languages the existence of lexical
and sentential iconicity can be called into question, this is not
the case across sign languages, given that iconicity is so abun-
dantly represented within any one language as well as across all
signed languages researched to date. Despite this fact, however,
early sign language research argued strongly against a role for
iconicity in language learning and language use. One initial force
driving the dismissal of iconicity was the attempt to disprove the
popular notion that sign languages were not true languages, but
merely gestural, pantomimic systems of communication. Since
arbitrariness of the sign is taken to be a basic tenet of language,
evidence was amassed that demonstrated that the iconicity
apparent in the surface form of signs was not important to the
language. Concerning language change, for example, Frishberg
(1975) concluded that iconicity erodes over time and iconic
forms become arbitrary. Although Klima and Bellugi (1979)
discuss the “coexistence of the iconic and arbitrary face of signs”
(p. 28) at length, they similarly stress that the iconic aspects of
signs are often inaccessible, in the sense that they cannot be
identified above chance by naive non-signers. Moreover, they
stress that signs consist of formational (phonological) features
that are themselves meaningless and that combine according to
a rule-governed system to make lexical signs in which iconicity
does not take part.

However, iconicity has not been completely ignored in sign lan-
guage research. Particularly in the realm of phonology, researchers
have recognized the potential role that iconicity plays in the for-
mation of signed languages. Mandel writes, “An adequate account
of (American) Sign Language must include the fact that the form
various elements take in the language depends in part on the visual
appearance of their referents” (Mandel, 1977, p. 57). In the litera-
ture, one can find thorough treatments of the rates and types of
iconicity found in various signed languages (Pizzuto and Volterra,
2000; Taub, 2001; Pietrandrea, 2002). Several phonological models
suggest that iconic properties of signs must be treated differently
than unmotivated aspects of form (Friedman, 1976; van der Kooij,
2002). Boyes-Braem (1981) proposed a lexical model that takes
semantically motivated elements of a sign into account. In her
analysis of ASL handshapes, she argued for an extra symbolic level
in addition to those levels present in spoken languages, which would
include iconically motivated handshapes, thus showing the impor-
tance of iconicity in the formation of signs. Friedman (1976), in
looking at sign locations (again in ASL), also suggested that iconic
locations be given special status. Under her analysis, iconic loca-
tions were treated as allophonic variations of a single phonemic
location. The idea of treating iconic features of a sign as existing
outside the realm of the phonology is also taken up in van der Kooij
(2002), who, in an attempt to constrain the phonology for Sign
Language of the Netherlands, determined that any form element
that makes a lexical contrast simply because it carries meaning
should be excluded from the phonological system and viewed rather
as a semantically motivated phonetic realization of a phonological

FIGURE 2 | Different forms of iconically motivated aspectual morphemes both singly and in combination in ASL (from Klima and Bellugi, 1979, permission pending).

FIGURE 3 | The signs meaning cat in ASL (A) and BSL (B).

<headnote>www.frontiersin.org December 2010 | Volume 1 | Article 227 | 5</headnote>
object. Under these accounts, iconic features of a sign are seen to have a more direct and presumably different link to meaning than unmotivated aspects of form.

Along similar lines, Semantic Phonology Theory (Stokoe, 1991; Armstrong et al., 1995) proposes that the phonological composition of a sign is meaningful, such that a sign’s meaning is partially predicted by its phonology and vice versa. Here, meaning is reflected in the phonological parameters of signs (hand configuration, movement, and place of articulation of signs) because they are grouped in and shaped by our cognitive and perceptual experience. Though this theory challenges the idea of the “duality of patterning” considered to be essential to language, there is clear evidence for separate levels of semantic and phonological information (i.e., a tip-of-the-finger state akin to tip of the tongue in which there is access to semantic information but not phonological information, Thompson et al., 2005). The role of iconicity has further received attention for its potential in shaping language structure and discourse in signed languages (Cuxac, 1999; Taub, 2001; Sallandre and Cuxac, 2002; Wilcox, 2004). Nonetheless, in terms of processing or language acquisition, the role of iconicity has been strongly argued against, based on what might appear to be very limited evidence, as we will discuss below in Section “Sensitivity in Signed Languages.”

ARE WE SENSITIVE TO ICONICITY?

SENSITIVITY IN SPOKEN LANGUAGES

Until recently, iconic form-meaning mappings have been considered uninformative about our understanding of what “language” is and, therefore, not worth serious consideration and empirical investigation. As discussed above, this viewpoint comes from both an historical-typological perspective (i.e., maintaining the idea of arbitrariness as the only fundamental feature of language, resulting in part from a focus on languages with very sparse iconic form inventories) and a processing stance which assumes that iconic mappings are so sparse in spoken languages that the system would not likely make use of them during processing. However, and despite this, there remains a rich history of investigation into the occurrence and generality of iconic form-meaning mappings. Some of this investigation stems from a desire to understand the origins of language, with iconic form-meaning mappings considered to be one possible and logical entry into the language system (Armstrong, 1983).

In line with this, Köhler (1929) was the first to show that speakers tend to match certain speech sounds to certain shapes. He found that speakers of Spanish reliably judged takte to be best associated with a jagged-edge shape and balaba to be best associated with a round, curvy shape. Similar results have been reported with speakers of different languages. For example, in a more recent study, Ramachandran and Hubbard (2001) asked monolingual speakers of English and Tamil “Which of these shapes is bobu and which is kiki?” Across both language groups, approximately 95% selected the curvy shape as bobu and the jagged one as kiki. These findings suggest that judgments are not tied to a particular language, but rather reflect sensitivity that is more universal in nature. Ramachandran and Hubbard (2001) argue that the “kiki/bobu effect” has implications for the evolution of language, since the naming of objects is not necessarily arbitrary. They point out a physical relationship between the shape of the mouth and the shape of objects being described. The rounded shape is more commonly labeled bobu likely because of the corresponding roundness of the mouth that occurs when producing the word. For the sharp, pointed kiki shape a matching tense mouth pattern produces the /i/ sound, and additionally the /k/ is harder and more forceful than a /b/. Ramachandran and Hubbard suggest that the human brain creates these cross-modal mappings (i.e., an initial rounded mouth representing a rounded visual shape gets linked to particular sound pattern) in a way akin to synaesthetics (people for whom stimulation of one sensory modality leads automatically to the evocation of experience in a second modality, as in the case of consistently seeing a number with a specific color), and that this cross-modal priming occurs through cortical connections among proximal cortical areas.

Along similar lines, other experiments have demonstrated participants’ ability to label objects at an above chance level using foreign (unknown) words (e.g., for South Malaita, Kiwai, Tongan, Finnish, and English: Gebels, 1969; Hebrew and English: Brackbill and Little, 1957; Japanese and English: Imai et al., 2008). In one study, for example, Brown et al. (1955) translated English antonym pairs into Chinese, Czech, and Hindi and asked English speakers to match them to the English translations provided. Not only were subjects above chance in making correct matches, there was also high agreement within the speaker groups. This is consistent with Sapir (1929) finding that English speakers show high levels of agreement in making comparative judgments between non-word pairs, e.g., in judging mal to be larger than mil. Brown et al. likewise speculate that speech may have originated from imitative connections between sounds and meaning, thus explaining their apparent universality. More recently, Imai et al. (2008) found (in the norming phase of an experiment discussed below) that both Japanese and English speakers made similar judgments on novel, but possible Japanese verbs, when judging whether or not they were iconic of particular actions. This finding emerged despite Japanese-speaking subjects’ knowledge of mimetic words in their language and English-speaking subjects’ ignorance of it. These findings lead Imai et al. to suggest that certain aspects of sound-symbolism are universally and biologically grounded. Similarly, Nygaard et al. (2009) found that sound-symbolic mappings facilitate word learning and processing cross-linguistically. Three groups of English speakers learned to associate Japanese words with an English translation word. One group learned the correct English translation equivalent, another group learned the translation equivalent of the Japanese word’s antonym, and the last group learned an unrelated translation word. Participants were subsequently asked to listen to the Japanese words and to pick the translation they had learned from two visually presented English words. The authors found that English speakers were faster to respond to Japanese words that had been learned with the correct English translation equivalent than those that had been paired with an unrelated English word. Moreover, the authors found a slight processing advantage for Japanese words paired with an antonym translation equivalent, which indicates that the effect may hold within conceptual/semantic domains. These findings suggest that sound-symbolic mappings are not arbitrary and language-specific, but rather reflect some more general phenomenon which extends cross-linguistically.
Iconic mappings in spoken languages further seem to be more resistant to regular sound changes. Joseph (1979) gives the example of the onomatopoetic interpretation of a cuckoo’s call, which diachronically has related forms, or cognates, in several Indo-European languages, all of which resemble the English word, cuckoo (i.e., Ancient Greek, Latin, Sanskrit and English). These cognates of cuckoo all failed to undergo an aspect of Grimm’s Law: namely the shift of voiceless stops to voiceless fricatives. The failure of these words to shift into something like /huhu/ is likely due to the closer link between the /k/ sound and the actual sound a cuckoo bird makes. Thus, there is evidence that languages conspire to preserve iconic form-meaning mappings.

Overall, considering judgments, there is clear evidence that speakers are sensitive to iconic form-meaning mappings. However, judgments are a rather indirect measure of language use, being off-line and susceptible to metacognitive strategies. More critical evidence for humans’ sensitivity to iconic mapping must come from on-line studies of language processing that show that iconicity affects lexical processes.

To date, there have been only a few experiments addressing the potential impact of iconicity in spoken language processing. It would seem that in this area, particularly, widely held beliefs about arbitrariness as a defining feature of language, and iconicity, conversely, as a trivial and almost paralinguistic feature, have likely hindered investigation. Nevertheless, what work has been done is suggestive. The results of initial research using off-line judgment tasks (discussed above) are indicative of the idea that common patterns of sound-symbolism should predict language-processing consequences. Specifically, there may be processing benefits (both in comprehension and production) for words that map more directly onto our perceptual and motor experiences of the world. There may also be benefits in terms of bootstrapping into the language system, a possibility discussed below.

Regular mappings between form and meaning have been shown to have a processing consequence in lexical decision tasks. Specifically, Bergen (2004) used lexical priming to investigate the role of regular form-meaning correspondences found for phonemes (e.g., /bl/ as in glint, glitter, glow, glare). Bergen found that these mappings sped lexical access over and above pure phonological or semantic priming alone, indicating a processing advantage for these form-meaning mappings. However, one must be careful in the case of phonemes as it is not completely clear whether the processing benefit derives from regularity or iconicity of form-meaning mappings. Specifically, priming from regular repeated phonemes like /gl/ and /owl/ may simply be a form of repetition-based priming. However, while clearly less representational, there is some evidence that speakers are sensitive to iconic aspects of these form-meaning mappings (see Bolinger, 1950; Jakobson and Waugh, 1979).

Of more relevance to our discussion are studies showing a processing effect between more clearly meaningful mappings. Westbury (2005) manipulated words (with either stops or continuants) presented in frames (black backgrounds with a white angular or rounded shape cut out of the middle). Subjects saw a string of letters, or a single letter/number in the middle of the frame and were asked to decide as quickly and accurately as possible whether a letter string was a word in Experiment 1, or, a single character was a letter in Experiment 2. Continuants (e.g., /ml/, /nl/) were recognized faster in a curvy frame, while stops (e.g., /p/, /k/) were recognized faster in a spiky frame, but the finding only held for non-words in the first experiment. Experiment 2 showed that the word-frame matching effect in Experiment 1 could not be due to shape-matching of the letters and the cut-out shape as there was no interaction of letter shape and frame shape.

If iconic mappings are common across languages and cultures, there may be some basic predisposition to mapping properties of visual objects and actions in the environment to specific acoustic properties. Shintel et al. (2006) looked at analog acoustic expression or “spoken gesture” to convey meaning related to direction of motion and speed of motion. Speakers describing the direction of motion of a dot on a computer screen reliably used a higher pitch for an upward moving dot (“It is going up”) than for a downward moving dot (“It is going down”). A significant difference in pitch was also found when participants simply read the descriptive sentences aloud, when presented on a screen. In a second experiment, the authors investigated whether the speed of object motion correlates with the speaking rate (an analog acoustic modulation). Speakers described dots moving horizontally across a screen as “It is going left” or “It is going right.” Crucially, the dots moved at different speeds (fast or slow), but speakers were instructed to describe only the direction of motion. Another group of participants was asked to listen to the recorded sentences and to judge the speed of the dots described in them. The results showed that both speakers and listeners used speaking rate to convey/comprehend information about an event, independent of the semantics of the lexical items. Speakers used a faster speaking rate to describe fast-moving dots than to describe slow-moving ones; and listeners could reliably guess the speed of the dot being described (although the descriptions themselves only encoded the direction of motion).

Thus, although limited in number, there is indeed some clear indication that iconicity affects spoken language processing. Showing that adult language users are sensitive to iconic form-meaning mappings begs the question of how these effects come about in development, and whether or not iconicity helps vocabulary learning.

Along similar lines to the Shintel et al. study, Walker et al. (2010) examined preferential looking patterns for infants and found that infants looked longer at a changing visual display (an animated bouncing ball) when it was accompanied by a sound (a sliding whistle) with a congruent pitch (high pitch for high location changing to low pitch for low location), when compared to incongruent pitch and spatial location. Thus Walker et al, provide evidence that even 3- to 4-month-old infants are sensitive to iconic mappings, thus addressing the question of whether or not these mappings must be learned or not. This finding suggests that these visual-auditory mappings are an unlearned aspect of perceptual cognition that may be the basis of form-meaning correspondences in language.

The role of prosody or intonational contours in early language development has been linked to characteristic vocalization patterns in infant directed speech (IDS), through which caregivers regularly...
convey specific communicative intentions like approval, praise, or warning (Fernald, 1989). In adult speech, prosody has traditionally been assumed to play a role in structural parsing, as a conduit for affective content and other aspects of communicative intention, but crucially, has not been assumed to be linked to word meaning in any way. However, Nygaard et al. (2009) have shown that speakers use prosody to process word meaning, and that certain concepto-semantc domains may have specific prosodic correlates. Participants heard novel words recorded with IDS-like prosody corresponding to different semantic dimensions (e.g., happy/sad, big/small, hot/cold). When shown pictures depicting each pole of antonym pairs (e.g., a happy person and a sad person) and asked to pick the one referred to by the novel word (e.g., “Can you get the seebow one?”), participants were able to do so by relying on domain-specific prosodic correlates to meaning. Importantly, when IDS-like prosodic contours were mismatched with semantic dimension (e.g., prosody for happy/sad occurring with pictures depicting hot/cold) participants were significantly poorer at picking the correct picture. This indicates that participants were not simply using prosody to cue valence, but rather that prosody does in fact convey meaningful information linked to specific semantic domains.

In a study more directly linked to iconicity encoded in the lexicon, Kovic et al. (2010) used both behavioral and electrophysiological measures to show that iconicity affects learning in adults. Kovic et al. used an implicit learning categorization task in which subjects learnt word/picture mappings implicitly (by making guesses and receiving feedback) and were then tested on these mappings. The pictures consisted of animal-like figures whose four prominent features (head, tail, legs, wings) were variously either rounded or pointed, and the words (i.e., the figure labels) were chosen according to sound-symbolic mappings with the figures (mot for roundness and riff for pointedness). Participants, assigned to either a congruent (i.e., round figures with mot) or incongruent condition (i.e., round figures with riff) in a training phase, were subsequently (in the testing phase) faster to confirm and slower to reject sound-symbolically congruent label-figure associations. Moreover, using event related potentials, an early negativity (N-200) was found for iconic mappings between the object and the label it was given when compared to mappings that were not iconic. Along the same lines as Ramachandran and Hubbard (2001), Kovic et al. (2010) conclude that the sensitivity to sound-symbolic (iconic) label-figure associations may reflect a more general process of auditory–visual feature integration where properties of auditory stimuli facilitate a mapping to specific visual features.

Specific to children, Maurer et al. (2006) found that 2.5-year-olds were sensitive to “kiki/bouba” correspondences. Children consistently matched words with rounded vowels to round shapes and words with unrounded vowels to the pointed shapes more frequently than the other way around and there was no difference between children performing this task when compared to adults performing the same task. That even young children are sensitive to these mappings in a language context further suggests their role in language development. Some aspects of iconic mappings between concepts and phonology have indeed been linked to facilitated learning. Imai et al. (2008) created novel verbs that were iconic (sound-symbolic) of particular actions, and other novel verbs that were not. They then used these novel verbs in a learning task with 3-year-old Japanese-speaking children. The children showed an advantage in learning the novel sound-symbolic (iconic) words, compared to the novel words that were not sound-symbolic, suggesting that regular mappings facilitate early language development. In line with this idea, it has been shown that Japanese children tend to learn iconic words very early on (Maea and Maeda, 1983).

Below we move to a discussion of studies addressing sign languages beginning with studies using off-line measures of subjects’ sensitivity to iconic mapping and then moving to studies using on-line measures and finishing with studies assessing language acquisition.

SENSITIVITY IN SIGNED LANGUAGES

In terms of sign language research, it might be surprising that despite the pervasiveness of iconicity at the sign (word) level there is almost no research investigating the degree to which signers are sensitive to iconic properties of a sign. Signers are clearly aware of iconicity, making use of it in areas such as poetry and word play (Sutton-Spence, 2005; this is also seen in spoken languages; see the book series: Iconicity in Language and Literature volumes 1–9, for a multitude of examples). Further, iconicity ratings have been collected for several sign languages (e.g., for BSL: Vinson et al., 2008; for ASL: Griffith et al., 1981; for ASL and DGS, Adam et al., in press), suggesting that signers are aware of iconicity in their language and are able to make judgments about it.

In a first study investigating possible consequences of iconic mappings in BSL, Vigliocco et al. (2005), showed an effect of iconicity in a similarity judgment task. They found that native BSL signers and English speakers differed in their judgments when grouping signs/words according to meaning similarity. In BSL, signs referring to tools (e.g., knife) and tool-actions (e.g., to cut) share “tool-use” iconicity. Vigliocco et al. (2005) found that, while English speakers tended to group tool-actions along with body actions (e.g., to hit), thereby preferring to distinguish actions from objects (and preserving a grammatical distinction between nouns and verbs), BSL signers tended to group tools and tool-actions together, as predicted on the basis of shared iconic properties of the signs. Vigliocco et al. account for the findings in terms of the mental images triggered by the iconic signs. In support of this, when English-speaking non-signers were instructed to create a mental image evoking typical experiences with the object or action, they behaved like the signers, judging tool-actions to be more similar to iconic tool-use (knife) than non-action signs (knife) and that body actions (e.g., to hit).

Our recent research has provided the first evidence for a clear role of iconicity in on-line language processing. In a first study, using picture/sign matching with ASL, Thompson et al. (2009) found that strong relationships between iconic properties of a...
sign and features of a pictured object speeded sign recognition
for signers. Specifically, participants were asked to indicate by
button-press whether a picture and a sign referred to the same
object. Experimental signs were all iconic. In one condition, the
iconic property/feature of the sign (e.g., BIRD, produced with
thumb and forefinger at the mouth, representing a bird’s beak)
was salient in the picture (e.g., a bird pictured from the front
with the beak well in view) while in the second condition the
iconic property was not salient (e.g., a picture of a bird flying
with the extended wings well in view). As a control, English-
speaking non-signers were also presented with the same pictures
followed by English words. ASL signers responded faster when the
iconic property of the sign was salient in the picture than when
it was not, while English controls showed no difference between
conditions. This first processing study provides evidence that a
more transparent mapping between meaning and form can aid
language processing.

In a replication of the ASL study, Vinson et al. (submitted) like-
wise found that BSL signers’ responses in picture-sign matching
were faster when the iconic property of a sign appeared saliently
in the picture, while non-signers showed no benefit in matching
such pictures to English words (replicating Thompson et al., 2009
in a historically unrelated sign language). The Vinson et al. study
further considered whether or not the “iconicity effect” could be
driven by typicality of iconic properties encoded in any one sign,
i.e., seeing a highly typical feature of an object such as the antlers
of a deer might help signers access meaning more quickly. Based
on English speaker-generated feature norms (McRae et al., 2005),
Vinson et al. controlled for typicality of iconic features encoded in
a sign. Specifically, the iconic aspect of half the signs represented
a highly typical feature according to the McRae et al. norms (e.g.,
whiskers for cat) and the other half represented an atypical feature
(e.g., pedals for bicycle; see Figure 4). The iconicity effect, i.e., faster
responses and higher accuracy for more iconic signs compared
to less iconic signs, was not modulated by feature typicality. No
difference in response times was found for signs whose iconicity
highlighted a salient, or more typical aspect of its referent compared
to signs that made use of less salient, but nonetheless iconic features
of a sign. The results indicate that the effect observed was general
to all iconic signs referring to objects, rather than stemming from
the typicality of the iconic property.

These findings establish the potential role of iconicity in lan-
guage processing, but leave unanswered the question of how these
effects come about. One could argue, for example, that they are
strategically linked to the task. Namely, subjects would not auto-
matically use iconicity in processing, although they would use it
in our experiments because the task specifically required them to
focus their attention on meaning-based relationships. The potential
use of metacognitive, task-related strategies that favor iconicity
effects were directly addressed in a further study in which we used
a phonological decision task, thus a task that does not require sub-
jects to focus on meaning properties. Specifically, Thompson et al.
(2010), asked BSL signers and non-signers to indicate whether a sign
employed a straight or bent handshake (a task based on the phono-
logical parameter of handshake). Despite the fact that the task did
not require subjects to access meaning, the relative iconicity of the
signs affected signers’ responses. As a further control, non-signers
participated in the experiment, as judgments about handshape
(form-based judgments) were possible without knowledge of a
sign’s meaning. Importantly, iconicity effects persisted even when
non-linguistic characteristics of the signs (e.g., differences in hand-
shape complexity or sign production time) were taken into account
by factoring out non-signers’ performance. Interestingly, signers
proved to be slower and less accurate to respond to iconic signs
than to non-iconic signs. Thompson et al. (2010) suggest that the
interference in making handshake-based decisions for iconic signs
could stem from more automatic access to meaning for iconic than
non-iconic signs which could serve to impair phonological-level
decision making for iconic signs. Overall, the findings show that
the iconic aspects of meaning inherent in phonological features of the
sign are automatically accessed even when they are not necessary
for a task (and, in this case, actually hinder performance).

To summarize, just as for the spoken languages, there are only a few
studies to date that directly assess iconicity effects (using offline
and on-line measures). Crucially, however, they suggest that
iconic mappings in signed languages play a role in processing and
are not simply an artifact related to specific tasks.

In terms of language development for signed languages, the
results have been more mixed. In a recent study, Ormel et al.
(2009) show that iconicity affects sign recognition by children aged
10–12 years. In a picture/sign matching task similar to Thompson
et al. (2009), pictures and signs (Sign Language of the Netherlands)
were displayed simultaneously on a computer monitor and subjects
were asked to decide if the picture and the sign matched. Ormel
et al. (2009) found that responses were significantly faster (to “yes
the picture matches the sign”) for highly iconic signs than for less
iconic signs. However, this is not particularly surprising given that

![Figure 4](image_url)
by age 10, children already possess a nearly adult-like vocabulary. The real test is whether iconicity affects vocabulary learning for younger children. Orlansky and Bonvillian (1984) provided some evidence that children's earliest signs are not iconic and Meier (1982) argues that iconic signs are not less prone to errors (e.g., for iconically motivated agreement signs such as GIVE which move from source to goal). Thus, Meier et al. (2008) found that sign errors produced during a longitudinal study of four deaf children (from as early as 8 months and continuing to as late as 17 months) did not tend to be more iconic than the correct sign (as one might expect to see, if children were tuned into iconicity; Meier et al. 2008). However, several researchers have argued that there are effects of iconicity on child language acquisition. Brown (1980) showed that hearing children (average age 4) were better at memorizing iconic signs for objects when compared to memorizing non-iconic signs. Brown argues that sign languages are easier to learn than spoken languages because of the high degree of iconic form-meaning mappings for these basic objects. This finding is consistent with adult second language learners of a signed language who make use of iconic properties of signs for learning (Lieberth and Gamble, 1991; Campbell et al., 1992). More importantly, early acquisition of iconic forms has been found in studies examining deaf children learning a signed language natively, Slobin et al. (2003), examined the productions of deaf children with deaf parents and found that iconicity promotes the early (before age 3) emergence of meaningful handshape distinctions, such as those found in handling classifiers (i.e., classifiers that depict an agent holding an object). Casey (2003) likewise concludes that there are effects of iconicity on the acquisition of ASL. Casey analyzes the naturalistic (children interacting with their caregivers) longitudinal data of six deaf children learning ASL natively from deaf parents. Casey notes that directional verbs, verbs that mark agreement (by moving to locations in space associated with verbal arguments), emerge earlier (at ages 1:6 to 2:1) with verbs such as give, which map more directly onto real-world movement. Interestingly, Casey notes a developmental continuity between non-linguistic action gestures and the very similar-looking linguistic use of verb agreement.

Thus, the existing evidence on the role of iconicity for language acquisition does not provide decisive evidence for or against the early use of iconic form-meaning mappings. A first, and important, point to note here is that these studies have taken iconicity as a holistic concept. However, different kinds of iconic mappings may vary in learning difficulty. Iconicity in signs can represent the actions of a real-world referent (e.g., TO-HAMMER in BSL which is produced with a handshake used to hold a hammer, along with the back and forth movement of hammering; motor iconicity), the form of a referent (e.g., OWL in BSL indicates the owl's large eyes; perceptual iconicity), or both (e.g., AIRPLANE in BSL is produced with the pinky and thumb extended to indicate the form of an airplane, and moves at head level in a straight line to indicate the action). Additionally, iconic properties of signs can indicate functional interaction with a referent (e.g., typing hands for COMPUTER) or even interaction with an object whose relation to the referent requires quite considerable world knowledge (e.g., milking hands of a cow's udder for MILK). It is important for future research to tease apart these different types of iconicity and to assess their relative learning difficulty (see Meier et al., 2008 for a similar point).

Cognitive developmental may also play an important role in a child's ability to use iconicity in the acquisition of form-meaning pairings. Namy et al. (2004) tested English-speaking children's ability to learn action-based iconic and arbitrary gestural labels (e.g., a hammering gesture vs. a dropping gesture) for objects associated with certain actions (e.g., a hammer). Learning followed a U-shaped trajectory with respect to the acquisition of arbitrary labels. At 18 months and again at 4 years of age, children are equally good at learning iconic and arbitrary labels for objects. However, at 26 months, children continue to learn iconic labels, but exhibit very poor learning of arbitrary labels. The authors argue that this decline in arbitrary symbol learning reflects a change in children's expectations and awareness about the forms that object labels can take. During this period of reorganization—in which regularities in the language system overall are being discovered—children seem to temporarily zero in on iconic mappings as a mechanism by which form-meaning mappings may be created.

Tolar et al. (2008) tested the ability to interpret iconic signs (i.e., ASL signs) by hearing children (2.5-5 years old) who had no previous experience with sign language. Children were shown a sign and asked to pick the picture that matched the meaning from a set of four pictures. At about 3 years of age, children began to exhibit patterns of performance suggesting a cognitive shift whereby they started to become explicitly aware of iconicity as bridging between concept and form.

In contrast to the Namy et al. (2004) study, which used only action-like gestures, Tolar et al. (2008) included iconic signs reflecting the action associated with a referent as well as signs that were iconic of perceptual (i.e., static) features of referents, thereby providing some initial assessment of the role of motor vs. perceptual iconicity. They found that, across all ages, children were better at understanding signs depicting actions, compared to signs depicting perceptual features. As Tolar et al. point out, these results are further consistent with hearing toddlers’ gestures, which more frequently imitate actions done with objects than they depict perceptual qualities of the object (Acredolo and Goodwyn, 1998).

Thus, whereas there seems to be clear evidence for a role of iconicity in language processing, the literature leaves open two alternative hypotheses concerning the role of iconicity in language development. On one account, iconicity would come into the picture as mediated by cognitive development that would allow children to grasp the link between features of the phonology of their language and properties of referents in the world. Under this view, iconicity per se would not help initial vocabulary acquisition in spoken or signed language, but could boost the learning later on, once meta-linguistic knowledge starts to develop. A prediction from this account is that we should observe a boost in the learning of iconic signs from a certain age, possibly around 3 (as suggested by Tolar et al.). On an alternative account, iconicity in sign language may rely on basic sensori-motor associations, hence iconicity effects (at least those reflecting basic motor and sensory-motor associations) would not be mediated by cognitive development and could support initial language development. Other types of iconicity that require more cognitive appreciation would come to play a role later. These two alternatives are explored in more detail below.
GENERAL DISCUSSION
To summarize our discussion in the previous sections, we have first shown that iconicity is a far more pervasive property of human languages than currently recognized: it is found in both spoken and signed languages and across all levels of language (i.e., syntax, morphology, and phonology). In Section “Are we Sensitive to Iconicity?” we have shown that the few studies which have assessed processing consequences of iconicity have found that iconicity does in fact affect processing. Moreover, a number of developmental studies show an advantage in learning iconic rather than arbitrary form-meaning mappings. The pervasiveness of iconicity in languages, and its role in language development (which we readily admit is controversial) and processing provides a challenge to the received view of language as an arbitrary, symbolic system. As Westbury states it: “Sound symbolism [Iconicity] effects challenge theories of language that posit encapsulated language systems that are wholly independent of other cognitive or sensory functions. They also challenge theories of language that posit encapsulated linguistic subfunctions, by assuming that phonology, orthography, semantics, and syntax are all processed quite distinctly.” (Westbury, 2005, p. 16). We believe that this is the first important lesson we can take from our review.

For spoken languages, the dismissal of this pervasiveness may originate from the fact that language studies historically focused on iconically impoverished languages, such as Indo-European languages. A really extensive language documentation, of course, requires resources for long-distance travel and prolonged stays in the “field.” In addition, comprehensive and fully representative language documentation (certainly for signed, but also for spoken language) has been greatly facilitated, or even made possible, through recent technological advances in video recording and archiving. For sign languages, pervasiveness of iconic expression is the rule, and one can speculate that if research on language had started with sign languages, the picture would look quite different—iconicity, rather than arbitrariness, would be heralded as the fundamental feature of linguistic forms. As an additional crucial consideration, research has thus far focused primarily on language in both modalities as being unimodal (i.e., the vocal modality for spoken language and the manual modality for signed language), rather than viewing language as a complex multichannel communicative system. As a result, other iconic forms accompanying speech or sign are often ignored or considered to be largely epiphenomenal.

THE ROLE OF GESTURAL FORMS IN COMMUNICATION
If we consider language as a complex of linguistic-discrete and imagistic-analog (or gestural) elements (McNeill, 1992; Okrent, 2002), we can make an even stronger case for a clear role of iconicity. Even speakers of iconically poor languages make abundant use of iconic links in their co-speech gestures. Co-speech gestures are tightly integrated with speech both semantically and temporally and often iconically represent aspects of the meaning being linguistically conveyed in speech. It has been shown that gestures enhance comprehension through mutual interaction and automatic integration with information from the speech channel (Chu and Kita, 2008; Kelly et al., 2010). From an embodiment perspective, gestures have been argued to be manifestations of the simulations of action and perceptual imagery that are involved in language production and comprehension (Hostetter and Alibali, 2008, 2010). In addition, as we have seen, the speech signal itself may also be gradationally (or gesturally) modulated in order to convey information related to the meaning of objects and events. Through “spoken gesture” – or analog acoustic expression – visual-spatial information in the world (e.g., speed of motion or direction of motion) can be encoded acoustically as modulations of pitch and speaking rate (Shintel et al., 2006). Similarly, prosodic variations in pitch and amplitude can convey information related to specific semantic domains (e.g., big/small, hot/cold) (Nygaard et al., 2009).

Signed languages, in addition to the iconic potential of the manu-
ally produced signs themselves, also exhibit regular and systematic use of iconicity in the non-manual features that accompany many signs. Through the use of “mouth gestures,” together with modula-
tions of the face and eyes, signers can convey additional information about visual-spatial features of objects and events, as well as about affect and stance (Sandler, 2009). Visual-spatial information that can be expressed gesturally in this way includes size and shape attributes (e.g., puffed cheeks and rounded lips for roundness), relational meanings (e.g., tightly pressed together, elongated lips and squinted eyes for narrow), and information about manner and path of motion (e.g., repeated puffing of cheeks and parting of lips for bumpy). Sign languages display a further correspondence between the mouth and the hands in what is called “echo phonology” (Woll and Sieratzki, 1998). Here, certain properties of the manual move-
ments of signs are reflected, or echoed, in oral components. For example, a separation of the hands or fingers may be accompanied by [pa], which involves a similar separation of the lips; while an oscillatory movement of the fingers or hands (as in the slight back and forth movement of the fists in the BSL meaning “not yet”) is accompanied by [shhh], an oral echo of the manual oscillation.

In this broader perspective, not only is the presence and preva-
lence of iconicity not surprising, but crucially it opens up new lines of investigation concerning how language builds upon and necessitates the involvement of other cognitive systems. To date, there exists a substantial literature on different aspects of iconicity (as our overview in Iconicity in Languages: Is it Really There? suggests). However, what has until very recently been missing is the recognition of iconicity as a basic principle guiding language use. Below we make an initial proposal of why it should be considered as such.

ICONICITY AS A FOUNDBATIONAL DIMENSION OF LANGUAGE
On the basis of the review above, we argue that any viable theory of language use must include iconicity in addition to arbitrariness as a guiding principle (at both the ontogenetic and phylogenetic level). We are not arguing, in fact, that we should completely abandon the idea of arbitrariness as a property of human language. This would obviously be incorrect, given the high frequency of arbitrariness in lexical representation, and indeed, its preponderance as compared to iconic mappings in many (families of) spoken languages. Rather, we would like to argue that both iconicity and arbitrariness are general principles of language, which both rep-
resent adaptation of specific languages to two fundamental con-
straints driving the phylogenesis and ontogenesis of the language system. These two fundamental constraints are the need to ensure an effective linguistic signal and the need to link linguistic form to human experience.
The need to ensure an effective linguistic signal would favor
arbitrariness. At the lexical level, for example it has been argued that
arbitrariness is an important design feature because it allows for
maximum discrimination between entries in a lexicon (Monaghan
and Christiansen, 2006) and thus allows for larger lexica to develop
(Gasser, 2004). As such, arbitrariness plays an important role in
terms of advantages to the language system, allowing a larger lexi-
con as well as increased communicative success (imagine how com-
municative effectiveness might be adversely affected if all things
within a particular semantic domain, e.g., tools, were phonologi-
cally similar). Along similar lines, Gasser (2004) and Haiman (1980,
1985) note that the degree of motivation (or iconicity) in a language
varies inversely with the size of its basic vocabulary. This is evident,
for example, in restricted registers (e.g., used in mourning or with
elders) and in pidgin languages. For example, in pidgin languages,
the relationship between antonym pairs is often morphologically
transparent (i.e., iconic) rather than opaque, as in New Guinea
Pidgin gutpefa (good) vs. no+ gutpefa (bad) (Haiman, 1985, p. 231).
In making this claim, Haiman pits the opposing forces of iconicity
and economy against each other. As frequency of use increases, the
need for economy and effectiveness of form drives down the use of
periphrastic (i.e., definition-like, and therefore more transparent)
expressions to denote concepts. This results in lexical elaboration,
that is, in an increase in vocabulary size, as well as in overall opacity
of lexical and grammatical contrasts. Like Haiman, we recognize the
interaction of different constraints. We are not denying here that
the need to ensure effectiveness of the linguistic signal drives the
system. However, we argue that this is not the only constraint.

The second constraint, namely the need to link linguistic form
to human experience, would favor iconicity. The importance of
linking linguistic form and human experience is central to embodied
views of language and cognition (and has long been advocated by
functionalist approaches to language, e.g., Bates and MacWhinney,
1982). According to embodiment, language comprehension, for
example, requires mentally re-enacting or simulating, the specific
embodied experience (e.g., Barsalou, 1999; Barsalou et al., 2003).
Despite ongoing debates concerning the extent to which this re-
enactment requires the same low-level systems used in perception
and action (see, e.g., Meteyard and Vigliocco, 2008 for a review),
there is strong evidence that some degree of embodiment is involved,
indicating that language use (i.e., production, comprehension, and
acquisition) requires that linguistic form activate the same systems
used in perception and action. Without such activation, communication
could not be successful. An important question then is how
words come to engage motor and perceptual systems.

At the onmogenetic level, and in the limited domain of lexical rep-
resentation, for words or signs with a purely arbitrary connection to
their meaning, one could argue that mere repeated temporal associa-
tion between a label (e.g., drink) and perception or production of the
corresponding action could ensure that the link between linguistic
form and sensori-motor experience becomes entrenched during
development (using a Hebbian-like mechanism, see Pulvermüller,
1999, or other more sophisticated co-activation type models Glenberg
and Gallese, under review). However, this may not be the whole story.
The pervasiveness of iconic mappings invites the hypothesis that
iconicity may also contribute to this process by providing scaffolding
for the cognitive system to connect linguistic form and embodied
experience. In other words, iconicity would provide a bridge across
the “major rift” between conceptual and linguistic form. It is an open,
empirical question whether this scaffolding is useful, or even nec-
essary, in jump-starting language development, or whether it only
comes into play after an initial set of form-meaning associations
are in place (developed, for example, on the basis of mere repeated
temporal associations). In either case, we would like to argue here that
the degree of interconnection and embodiment we observe in adult
language processing is made possible, or at least greatly facilitated,
by the existence of iconic form-meaning mappings.

Children and their caregivers make ample use of iconic, analog
representation in interactive and communicative situations. For spoken
languages the use of iconic forms is prevalent in the prosody of
IDS, in the use of onomatopoeia, and in the use of spoken and manual
gesture. All of these may bridge the gap between our experience of the
world and our ability to communicate about it – ultimately between
conceptual and linguistic form – because they are analog mappings
linking to analog referents in the world. The analog relation to refer-
ents helps ground linguistic forms into our perceptual-motor experi-
ences of the world, giving rise to embodied language (i.e., language
grounded in our sensori-motor experience). The idea of iconicity as
providing a way of conveying gradient, analog information is crucial
to our suggestion that iconicity is the scaffolding that allows our
language/communication system to “hook up” to our experience of
the world, and to ground this experience in our perceptual and
motor systems. We may think of iconic (analog) forms as providing
the necessary link to development and use of a more symbolic system,
which utilizes discrete, arbitrary (as well as iconic) forms. What we
perceive and experience in the world is non-discrete, continuous,
and analog in nature. Thus, the use of representational forms of the
same nature may be crucial in building up a communication system
that is grounded in our experience of the world.

Phylogenetically, we may speculate that some degree of iconicity
should be present in order to facilitate (or even render possible) the
mapping between linguistic form and meaning. Languages, nonethe-
less, can differ greatly synchronically with regards to the extent to
which iconicity is incorporated in phonological form, for example,
rather than in other aspects of communication (such as co-speech
gestures, as discussed below). The scope and significance of the large
variability across human languages in amount of iconicity is an open
question calling for further investigation. However, the claim we are
putting forward regarding the role of iconicity in language evolution
supposes that iconicity was an essential ingredient in the transforma-
tion of early forms of communicative interaction into the complex
language systems we master today. Our claim may resonate with
suggestions that gestural (i.e., manual gestural) communication was
an evolutionary precursor to vocal communication (Rizzolatti and
Arbib, 1998; Stokoe, 2001). However, in contrast to a “gesture-first”
evolutionary path into human language, we take the position that both
manual and vocal gesture would be necessary as a precursor to our
developed language systems. Articulations in both channels provide
iconic mappings that would allow the building up of a symbolic sys-
tem on the scaffolding of analog relations between representational
forms and the world. A “gesture-first” model of language evolution
is saddled with the burden of explaining the transition from manual to
vocal articulation. In fact, McNell (2005) argues that the assumption
that language started in the manual channel and then switched to the
vocal channel is incompatible with the evolution of language into a
complex communication system in which linguistic (e.g., speech)
and imagistic (e.g., co-speech gesture) components are tightly and inextricably integrated. That is, the tight integration of the two systems (linguistic and imagistic) synchronically suggests the connection has been in place from the beginning, and indeed was instrument in jump-starting the system (McNeill, 2005). The phenomenon of echo phonology in sign languages lends further support to this notion. Phylogenetically, the tight link between the manual and vocal articulatory systems provides a possible mechanism for the evolu tionary transformation of iconic forms into arbitrary forms (Woll and Sieratzki, 1998). Ontogenetically, the manual and vocal systems, which have access to iconic, analog expression in different ways and to differing degrees, and which thus exhibit direct links to the world of experience in different ways, are used in conjunction to build up a representational system for the world. These correspondences facili tate the grounding of experience with the cognitive system.

Thus, to conclude, our review makes it clear that iconicity is present across spoken and signed languages and that it plays a role during language processing and possibly in language acquisition.

We propose that iconicity is exploited to the service of guarantee ing the link between linguistic form and human experience. The variability in the forms and amount of iconicity across languages indicate different manners in which languages can get the balance right between two basic constraints, namely the need to link lan guage to our experience (which would favor iconicity) and the need to have an efficient communication system (which would favor arbitrariness).

ACKNOWLEDGMENTS

This work was supported by a Marie Curie Fellowship (MEIF CT-2006-041886) awarded to Pamela Perniss and the Economic and Social Research Council of Great Britain (Grant RES-620 28-6001) grant awarded to the Deafness, Cognition and Language Research Centre (DCAL). We would like to thank Neil Fox, Clifton Langdon, and Mark Nelson for their valuable help with figures as well as David Vinson for insightful comments on an earlier draft.

REFERENCES


Implementation and Iconicity, LOT, Utrecht.


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received 11 May 2010; accepted 01 December 2010; published online: xx December 2010.


This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

Copyright © 2010 Pernis, Thompson and Vigliocco. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.