PNAS

Main Manuscript for

Iconicity as an Organizing Principle of the Lexicon

Erin E. Campbell, Boston University
Zed Sehyr, San Diego State University
Elana Pontecorvo, Boston University
Ariel Cohen-Goldberg, Tufts University
Karen Emmorey, San Diego State University
Naomi Caselli, Boston University

Keywords: language; iconicity; systematicity, phonological similarity, semantic associations

This PDF file includes:

Main Text
Figures 1 to 5
Tables 1 to 4

Abstract

The view that words are arbitrary is a foundational assumption about language that has been used to set human languages apart from non-human communication. We present here a study of the alignment between the semantic and phonological structure (systematicity) of American Sign Language (ASL), and for comparison, two spoken languages—English and Spanish. Across all three languages, words that are semantically related are more likely to be phonologically related, highlighting systematic alignment between word form and word meaning. In ASL and English, though not in Spanish, there was a significant effect of iconicity (perceived form-meaning resemblance) on this alignment: words are most likely to be phonologically related when they are semantically related and iconic. This phenomenon is particularly widespread in ASL: half of the signs in the ASL lexicon are *iconically* related to other signs, i.e., there is a non-arbitrary relationship between form and meaning that is shared across signs. Taken together, the results reveal that iconicity can act as a driving force behind the alignment between the semantic and phonological structure of spoken and signed languages, but languages may differ in the extent that iconicity structures the lexicon. Theories of language must account for iconicity as a possible organizing principle of the lexicon.

Significance Statement

Contrary to the widely held assumption words are arbitrary symbols for their meaning, this study shows that when words have similar meanings, they are also similar in phonological form. Further, we find that iconicity (a physical resemblance between word form and word meaning) influences this alignment in American Sign Language and English, but not in Spanish. Iconic systematicity is present in half of all signs but rarer in the spoken languages studied. These results suggest that rather than a fringe phenomenon, motivated form-meaning associations can act as fundamental organizing characteristic of at least some human languages.

Main Text

Introduction

Until the 1970s, the field of Linguistics relied exclusively on investigations of spoken languages. As a consequence, linguistic properties that are pervasive in sign languages but limited in speech played a marginal role in linguistic theory. Perhaps the starkest example of this is iconicity, where the form of a word or sign can resemble its meaning. While sign languages have phonological structure (i.e., handshapes, locations, and movements are combined in rule-governed ways to create signs, just as speech sounds are combined to create words), signs can also resemble their meanings (e.g., the American Sign Language (ASL) sign STIR[1] bears a perceivable physical resemblance to the act of stirring). Although a substantial proportion of sign language lexicons appear to contain some degree of iconicity, words in spoken languages are thought to be largely arbitrary—their pronunciations bear little to no perceivable relationship to their meaning (e.g., words meaning 'to stir': *remuer* (French), *míchat* (Czech), *Khn* (Thai), etc). The historical bias for spoken language led to the assertion that arbitrariness is a defining feature of human language (Whitney, 1867; Saussure, 2001; Hockett, 1960). The central idea was that because the forms of words are not tied in any way to their meaning, language serves as an abstraction from direct perceptual experience and allows humans to communicate infinite meanings in infinite contexts using only a finite set of speech sounds. This claim that languages are fundamentally arbitrary still holds a central position in linguistic theorizing, and popular linguistics textbooks still make claims like: "Despite occasional iconic characteristics, human

[1] In this paper, English glosses of signs are written in uppercase. For the convenience of the reader, we have provided links to videos depicting these signs.

language is essentially arbitrary" (Finegan, 2014), and list arbitrariness as a defining property of human language (e.g., Yule, 2023).

Unfortunately, the centrality of arbitrariness to early linguistic theory led researchers to ignore or, at best, downplay the ways in which form and meaning are in fact related in human language. The existence of onomatopoeia (words whose phonetic form mimics their referent's sound, e.g., 'whoosh', 'meow', 'whack') and phonesthemes (sequences of sounds that semi-systematically occur in words with particular semantic features e.g., /sn/- is used in words relating to noses: 'snort', 'sniff', 'sneeze', 'snout') has long been recognized in spoken languages (e.g., Firth, 1930). However, despite their consistent cross-linguistic presence, these phenomena are often considered marginal exceptions to the rule: that words are arbitrary symbols.

In recent decades, however, research into the non-arbitrariness of language—both signed and spoken—has continued to grow. Across the literature '**arbitrariness'** is positioned as the opposite of two related but not identical constructs: iconicity and systematicity. We operationalize **'iconicity'** as a motivated, structured mapping between the perceptual properties of a person's phonological and semantic representation of a lexical item (Winter et al., 2023). For example, in ASL the motion and location of the two hands in the sign RECORD correspond to two reels in a cassette or "reel-to-reel" player[2]. '**Systematicity'** is when two or more lexical items share both meaning and some elements of phonology. For example, /gl/ appears across English words relating to light ('glimmer', 'gleam', 'glow'). In this paper, we explore both iconicity, systematicity, and the intersection of the two. Systematicity need not be iconic (e.g., /gl/ bears no resemblance to light), which we refer to as **non-iconic systematicity**. In some cases, systematicity and iconicity co-occur (**iconic systematicity**), as in the cluster /sn/, (which includes a nasal sound and is used in words relating to noses: 'snort', 'sniff', 'sneeze', 'snout'; this particular mapping is common cross-linguistically; Blasi et al., 2016). See Dingemanse et al. (2015) for a detailed overview of the intersection of these phenomena. Studies have also demonstrated that iconicity plays an important role in language evolution, language acquisition, and language processing (see Perniss et al., 2010, for review). Sign languages in particular exhibit sophisticated mechanisms of iconic form-meaning mapping, and recent work has demonstrated important cross-linguistic generalizations for how meaning may be expressed in sign-based phonology (e.g., Börstell et al., 2016; Hwang et al., 2016; Östling et al., 2018; Padden et al., 2013). Collectively, the pervasiveness of these findings forcefully indicates that iconicity must be the subject of linguistic inquiry.

Figure 1.

While much of what we know about iconicity has focused on the mapping between a single lexical form and a single meaning, a handful of studies have identified iconic mappings that occur systematically across the lexicon. These studies have generally proceeded in a top-down

[2] Iconicity is inherently subjective, and depends on a person's perceived relationship between semantic and phonological representations (i.e., RECORD might not be iconic to a person who has only ever used it to refer to electronic recordings).

fashion where researchers identified a potential iconic vehicle and then collected data to adjudicate whether this relation was in fact borne out within or across languages. For example, Börstell et al. (2016) observed that signs for concepts that inherently include multiple entities (e.g., eyes, shoes, family) are frequently denoted by two-handed signs. Similarly, Occhino (2017) examined signs in ASL and Libras (Brazilian Sign Language) that used a claw handshape, and found that across both languages convex/concave shapes were frequently denoted by signs using a claw handshape (such as BALL or BOWL in ASL). Studies of spoken languages also often examine researcher-identified associations between certain phonemes and meaning (e.g., /k/ with sharpness in the classic bouba/kiki effect; Köhler, 1929).

The top-down approach exemplified by Börstell et al. (2016) and Occhino (2017) allows researchers to rigorously collect evidence for the existence of individual iconic mechanisms and has revealed important high-level patterns across languages (e.g., Padden et al., 2013). However, a drawback of this approach is that it takes researcher-identified vehicles as its starting place and thus may miss subtler iconic form-meaning relationships in favor of ones that are highly salient or more readily perceivable. Second, this approach sometimes involves focusing on selected sets of signs (e.g., those that are conceptually plural), making it difficult to identify, describe, or quantify broader patterns across the lexicon.

Lexicon-wide analyses have revealed that there is more systematicity than expected by chance in both spoken and sign languages (Monaghan et al., 2014; Martinez del Rio et al., 2022). Martinez del Rio et al. (2022) documented both iconically-motivated systematic relationships between iconic signs in ASL (thinking signs like KNOW, THINK, and MEMORIZE are often produced on the head) and non-iconically-motivated systematic relationships (e.g. familial relationships like MOTHER, AUNT, COUSIN are often produced on the head). However, it is unclear whether iconic systematicity is restricted to small pockets of the ASL lexicon, or whether iconic systematicity is a pervasive pattern of form-meaning correspondence that exists across languages.

In the present project, we used a bottom-up approach to explore iconic and non-iconic systematicity across the lexicon of ASL, and for comparison, spoken English and Spanish. As above, systematicity is defined as systematic alignment between words' phonological and semantic properties, i.e., similarity in meaning is systematically associated with similarity in form. We determined semantic similarity using data from semantic association tasks (described more extensively in Methods), in which native users of each language were asked to provide the first three words that came to mind in response to a cue word. We then integrated semantic similarity with estimates of phonological similarity between the pairs. This approach enabled us to identify alignment between semantic and phonological relatedness in pairs of words in each language (using three large datasets). With these data, we asked: In each of the languages, does iconicity magnify the relationship between semantic and phonological similarity such that word pairs are most likely to be both phonologically and semantically related when they are iconic?

Results

Statistical Analysis

We conducted a mixed effects linear regression analysis that predicted the phonological distance between pairs of signs in ASL (measured in number of differing *phonological features*), based on the items' iconicity rating and semantic relatedness. The model included as predictors: semantic relatedness (related/a semantic associate, or unrelated/not an associate), summed iconicity ratings for the pair, and an interaction between semantic relatedness and iconicity. The models also included random intercepts of each word in the pair. We found a significant main effect of semantic relatedness, such that semantically related sign pairs tended to be more phonologically similar than semantically unrelated pairs (β= -1.10 [-1.14 – -1.06], p < 0.001). The model also contained a significant main effect of iconicity (β= -0.06 [-0.09 – -0.03], p < 0.001), wherein highly iconic pairs were more phonologically similar than less iconic pairs. Critically, there was also a significant interaction between semantic relatedness and iconicity (β= -0.13 [-0.16 – -0.09], p < 0.001): Specifically, iconicity magnified the relationship between semantic associations and phonological similarity; the difference in phonological similarity between semantically related and unrelated signs was *larger* when the signs in the pair were more iconic (see Figure 2.A).

One might reasonably wonder whether this effect of iconicity on form-meaning relationship is due to ASL being a visuo-gestural language: could iconicity play a different role in the lexica of signed vs. spoken languages? We therefore fit two parallel models onto spoken languages, one for English and one for Spanish, where we predicted the phonological similarity between pairs of words (measured in number of differing *phonemes*) based on semantic similarity, iconicity, and an interaction between semantic relatedness and iconicity. Parallel to the ASL results, we found that semantically related word pairs tended to be more phonologically similar than semantically unrelated pairs (English: β= -0.16 [-0.16 – -0.16], $p < 0.001$; Spanish: β= -0.32 [-0.33 – -0.30], $p < 0.001$), and highly iconic pairs were more phonologically similar than less iconic pairs (English: β= -0.22 [-0.24 – -0.20], $p < 0.001$; Spanish: β= -0.12 [-0.18 – -0.06], $p < 0.001$). In English, as in ASL, we additionally observed an interaction between iconicity and semantic relatedness wherein iconicity seemed to amplify the effect of semantic relatedness on phonological relatedness (β= -0.03 [-0.04– -0.03], $p < 0.001$; Figure 2.B). In Spanish, however, there was no significant interaction (β= 0.01 [-0.01 – 0.02], $p = 0.340$): the effects of semantic relatedness and iconicity were additive, rather than interacting (Figure 2.C). Summaries of model results are provided in Figure 2 and Tables 1—3.

Data Visualization

In order to explore phonological and semantic relationships between word pairs, we next present network visualizations of systematicity for each language. In these networks (Figures 2-4), each node represents a word. We first filtered each dataset to wordpairs that were systematically related to each other, as defined by: 1) they shared at least half of their phonemes with each other and were freely associated by at least 2 participants—wordpairs that met these criteria were connected by an edge. We then excluded words that did not share at least one edge with another word. Note that this definition of semantic relatedness is stricter than in the statistical analysis (e.g., we require that at least two versus one person associated the items, excluding words that were not related to any other words)—this is because more aggressive pruning helped to simplify the network visualization. While the pruned network is more interpretable, it is based on an arbitrary cut-off for whether pairs are systematically related, and thus does not include the full continuum of possible systematic relationships. We then organized the networks using the Fruchterman-Reingold algorithm, which is a force-directed algorithm that attempts to identify an equilibrium in a network such that nodes are generally repelled from one another, but nodes that share an edge are drawn toward one another. For each network, we measured the percentage of words that were connected to at least one other word, the average number of connections for each node (mean degree), and the percentage of nodes in the network that formed part of the largest connected component (giant component). Within the giant component, we also measured the shortest path between each word and each other word in the component (mean path length) and whether a word's associates shared direct connections with one another (clustering coefficient). To view these networks, see Figures 2—4, and see Table 2 for network measurements. Additionally, to examine the network structures of high- and low-iconicity words, we filtered the nodes in the networks using a within-language median split on iconicity. We then re-computed network statistics for the high-iconicity and low-iconicity words. For these filtered networks, see figures S1-S6 in SI

What we observe in the network plots is that the ASL lexicon is rife with systematicity. Based on our criteria for systematicity (word pairs semantically associated by at least two people and share at least half of their phonological features), 76% of the ASL words ($n_{systematic/total}$=2,081/2,723 words) were included in the ASL systematicity network. The ASL network was characterized by relatively long chains of sparsely connected signs (Mean path length = 16.5; 34.8% of nodes appeared in the giant component), and signs tended to be systematically related to multiple other signs (Mean degree = 1.42).

Visualizing the data in this way showcases clusters of signs that share elements of form and meaning. For example, the ASL systematicity network (Figure 3) includes a cluster of iconic signs that related to food and are produced near the mouth (e.g., EAT, DRINK, SPICY, as described in Hwang et al., 2016) and a cluster of signs that relate to family and are produced on the head (e.g., as expected, based on Martinez del Rio et al. 2022: FATHER, MOTHER, GRANDFATHER).

That our network plots highlight known clusters of systematically-related signs lends confidence to our visualization approach. Our networks also identified neighborhoods of systematically-related signs that, to our knowledge, have not been previously described in the literature: including a group of non-iconic signs related to emptiness that use the same middle-finger-extended handshape (e.g., EMPTY, INVISIBLE, VANISH, NAKED), iconic signs that relate to weather produced with two hands in neutral space with all fingers selected (e.g., WIND, CLOUD, RAIN), and iconic signs that relate to exercise are produced with two closed hands (e.g., GYM, EXERCISE, STRETCH). We note that the systematicity in the ASL network is often, but not always, iconic.

Turning to the spoken language networks, 39% of the English words ($n_{systematic/total}$=1,867/4,829 words; Figure 4), and 28% of the Spanish words ($n_{systematic/total}$=380/1,374 words; Figure 5) were determined to be systematically-related to other words and thus appeared as nodes in the respective networks. Descriptively, words in the English and Spanish systematicity networks tended to occur in pairs or triplets rather than long chains (English mean path length = 5.19, Spanish mean path length = 1.61), and words in the spoken language networks tended to be connected to just one other word, if any (English mean degree: 0.832, Spanish mean degree: 0.742). Still, the English network highlights a chain of iconic words related to short, well-defined sounds (bleep—beep—bop—pop), a cluster of words for vocalization of distress (squeal—squeak—squawk), and a cluster of words related to bouncy, back-and-forth movement (wiggle—jiggle—giggle). Also visible are several previously described phonesthemes (the /sn/ in sniff—snuff, the /sh/ in mush—mash—smash; Hutchins, 1998). In Spanish, where we found no significant interaction between systematicity and iconicity, we observed that most word pairs or clusters could be attributed to morphology (organizar—organizacion; mirar—mirado) or etymology (cristo—cristiano; junio—julio), a point we return to in the Discussion.

Discussion

In this study, we took a data-driven, bottom-up approach to measure iconicity as a possible driver of systematicity in the lexicons of ASL, English, and Spanish. In order to uncover patterns of systematicity—neighborhoods of words in the lexicon where words are phonologically and semantically related—we combined information about the semantic structure of the lexicon (from semantic free association tasks) with information about their phonological structure (shared phonological features). Replicating earlier work, we found evidence for systematicity across all three languages: words that are closer in meaning tend to have more similar word forms. We then explored iconicity among words that were both phonologically and semantically related. This bottom-up approach to identifying iconic systematicity freed us from the need to rely on researcher-identified iconic motivations. In all three languages, iconicity predicted phonological similarity, suggesting that iconic motivations might be best expressed with some common phonological features. Crucially, we observed that in ASL and English, but not in Spanish, there was a significant interaction such that the relationship between semantic and phonological similarity was magnified for iconic pairs. Taken together, this work shows that iconicity is not simply a rudimentary use of the body to mimic the meaning of a single word, rather iconicity occurs in highly patterned ways across the lexicon. These results implicate form-meaning systematicity, and (in some languages) iconicity, as defining characteristics of the lexicon. These findings thereby challenge the longstanding notion that arbitrariness is a design feature of language (Whitney, 1867; Saussure, 2001; Hockett, 1960).

The way we have defined form-meaning systematicity allows us to identify sub-lexical units of form and meaning that occur across lexical items, and one might reasonably wonder what the difference is, if any, between form-meaning systematicity and morphology. Indeed, many of the systematic relationships we see shared across the lexicon look something like morphology. For example, in the ASL network, compound signs (e.g., PARENTS and MOTHER were clustered together, and PARENTS is a compound sign with the constituents MOTHER and FATHER). The ASL network also identified noun-verb pairs in which the noun is a reduplicated form of the verb (e.g., OPEN_BOOK and BOOK, GATE_CLOSE and GATE) and initialized signs (meaning the signs are formed the same way except the handshape relates to a letter of the manual alphabet; e.g., M for MONDAY, T for TUESDAY etc.). Similarly, the English network contains several vowel-change verb pairs (e.g., drive—drove, take—took), and the Spanish data showed links between words that were derivationally related to each other (e.g., necesario—necesidad, importante—importancia). However, in other cases, the shared elements are not discrete and/or productive, e.g., DRAGON, SPICY (related by an iconic or metaphoric depiction of flames at the mouth), or FIANCE, RING (related by metonymy). In some cases, pairs encode syntactic information (e.g., the ASL noun-verb pairs), but in many cases they do not. Some theories of morphology account for these facts more readily as they do not require morphemes to have discreteness, productivity, or syntax (e.g., Jackendoff & Audring, 2020), whereas other theories of morphology (e.g., Lexeme-Morpheme Base Phonology; Beard & Volpe, 2005) may find more friction. And conversely, some descriptions of systematicity include morphology (alongside iconicity and etymology) as drivers of systematic form-meaning relationships (Haslett & Cai, 2023).

In the ASL data, we found that many ostensibly non-iconic signs do not bear an obvious physical resemblance to their meaning, but in the context of other signs with similar forms and meanings, the shared elements reveal possible iconic mappings. For example, in our networks, we observed a group of signs related to employment status and are produced with the index and middle finger extended and moving relative to a closed fist (JOIN, QUIT, RESIGN). While in isolation, the sign QUIT may not strongly iconically evoke the act of quitting, together, the signs JOIN, QUIT, and RESIGN could be seen as depicting a person with two legs moving toward or away from an entity or organization. These patterns point to an even *larger* role of iconicity in the lexicon than estimated by our present approach. Our sense is that most of the systematic relationships in ASL could be analyzed as being iconically-motivated.

Interestingly, the converse was true for English and Spanish. That is, there was a striking lack of iconic motivations shared between words that are systematically related. Instead, systematically related words often appear as etymologically-related pairs (e.g., junior—senior, brother—mother, hotel—motel, lengua-lenguaje, nada-nadie), or binomials that are frequently collocated in speech (e.g., flip—flop, nature—nurture). Additionally in the spoken languages the analysis revealed couplets of systematically related words, whereas in sign language systematicity appeared across larger neighborhoods of more interconnected signs. This lack of shared iconic motivation across systematically related signs was borne out statistically in a modest interaction between iconicity and semantic relatedness on phonological similarity in English and a lack of any interaction for Spanish. Instead, our modelling approach only highlighted relationships between iconicity and phonological form, showing that word pairs that are more iconic are more phonologically similar (e.g., bang—boom, whoosh—swish). This result might reflect that certain speech sounds are used more often than others to depict auditory phenomenon (e.g., the "sh" sound to depict friction through air; plosive sounds like "b" to depict explosion) (Varda & Marelli, 2022). We might not see the interaction in Spanish either due to the big difference in statistical power ($N_{English}$ = 26,820,266 wordpairs vs. $N_{Spanish}$ = 1,883,756 wordpairs) or due to true cross-linguistic variability in whether iconicity is a primary driver of systematic form-meaning relationships.

One reason the patterns may differ across languages in this study is the number of phonological features that contribute to iconicity. While onomatopoeias and ideophones—some of the most salient examples of spoken language iconicity—tend to make iconic use of many of a word's sounds (e.g., "splash"; Dingemanse, 2012), these words are relatively rare in most languages. In many cases it is often only a single phoneme or phonological feature that seems to participate in the iconic mapping (e.g., in "poke", the action appears to be depicted by the stop manner of the consonants, but not their voicing or place, nor anything about the vowel). Conversely, in ASL many features often participate simultaneously in the structure mapping (e.g., the handshape and movement of the hands in RECORD). Building off the structure mapping theories of Gentner (1982) and Emmorey (2014), analogies are stronger when there are multiple connections between the target domain (e.g., phonology) and the source domain (e.g., meaning). When only one aspect of a word's phonology participates in form-meaning mappings, as is commonly the case in spoken language, iconicity may not be able to support strong systematic mappings across multiple lexical items. We suggest that the strength of iconic structure mapping (i.e., the number of phonological features that can map to semantic features) is a useful property on which to structure lexical relationships. It may also be harder to detect examples of iconic systematicity in languages where the form-meaning unit only makes up a small portion of the word. Another source of variability in iconic systematicity may be the degree of iconicity: It subjectively appears that the “highly-iconic” signs in ASL are more iconic than the “highly-iconic” words in the spoken languages, but it is not readily clear how to operationalize cross-linguistic differences in lexical iconicity (but c.f. Thompson et al., 2020). Linguistic theories may need to accommodate variation in the prevalence of iconic systematicity by considering other ways that systematicity may be instantiated (e.g., via morphology or etymology).

The multiply-layered relationships inherent to *iconic* systematicity, in particular, has consequences for psycholinguistic theories of how words are organized and retrieved from the mental lexicon. To date, psycholinguistic studies of iconicity in recognition and production have largely focused on the iconicity of individual signs (e.g., Bosworth & Emmorey, 2010; Caselli, Emmorey, & Cohen-Goldberg, 2021; Ortega & Morgan, 2015), what we refer to as ‘sign-specific iconicity.’ These studies ask questions about the extent to which signers are sensitive to iconicity in sign perception and production, for example. The effects of sign-specific iconicity on sign production and recognition have been quite mixed, and it is yet unknown whether and how lexicon-wide patterns of iconicity affect language processing. For instance, spreading activation has long been considered a factor in how individuals access and retrieve words in semantic memory (Collins & Loftus, 1975), whereby during word retrieval, other closely-connected words in the network also become active. Much of the spreading activation literature focuses on semantic and phonological organization of the network, but the role of iconicity has not been accounted for in possible patterns of spreading activation (e.g., are iconically-related lexical items more likely to become active during lexical access due to overlapping phonological *and* semantic properties?). Considering systematic iconicity in addition to or instead of sign-specific iconicity may help clarify the role of iconicity in sign processing.

Iconic systematicity may also play a role in how people learn words. There are robust effects of iconicity on early vocabulary acquisition in both spoken and signed languages (e.g., Thompson et al., 2012; Caselli & Pyers, 2017), but the mechanisms underpinning these effects remain elusive. While children might observe an analogy between the form of a new sign and its meaning (e.g., the perceptual similarity between trees and TREE) in ways that could make it easier to learn the sign, it is unclear whether this type of analogical reasoning can fully explain effects of iconicity that are detected during infancy (Gappmayr, Lieberman, Pyers, & Caselli, 2022). However, the prior studies of iconicity in word learning have focused exclusively on iconicity at the level of individual signs. It is possible that the effects of iconicity that have been documented are better attributed to iconic systematicity. By hypothesis, children might leverage iconic mappings between form and meaning in signs they *already* know to help them build new lexical representations of novel, iconically related signs. For example, a child might more easily learn a novel sign related to eating that is produced at the mouth (e.g., DEVOUR, TASTE) if they already know other eating-related signs that are also produced at the mouth (EAT). In English, effects of systematicity on words’ age of acquisition have recently been documented (Cassani, 2022): words whose phonology better conforms to English’s systematic form-meaning relationships tend to be learned earlier. Perhaps the effect of systematicity on age of acquisition is stronger in languages, such as ASL, that have more robust patterns of systematicity. How then might iconicity and systematicity interact in children’s acquisition of words, and does this change across languages that differ in the extent to which iconicity drives systematicity?

Limitations. The methods used to identify systematicity in this study represent just one out of many possible approaches. Semantic similarity might alternatively be measured by weighting according to the number of people who associated the pair, or with a different measure of semantic relatedness altogether (e.g., collocation in a corpus). Phonological similarity might be alternatively measured over a different set of features or by weighting features according to their position in a feature geometry. Perhaps because of how similarity was defined, the networks presented here may include patterns that others might not ascribe to systematicity (e.g., morphological relationships), and/or exclude real examples of systematicity. Signs may be related, but not meet our stringent criteria (e.g., eight shared phonological features, semantically associated by at least two people). In contrast, the methods used to identify iconic systematicity might have been overly inclusive, capturing sign pairs that are phonologically and semantically related and are iconic, but the iconic motivation for each member of the pair is unrelated (e.g., PULL depicts a person pulling a rope, and MAGNET depicts two small objects being drawn toward one another). Understanding how different operationalizations of phonological and semantic relatedness impact observed systematicity may further illuminate the structure of form and meaning correspondence in the lexicon.

Conclusion. In sum, we found evidence that challenges the idea that arbitrariness is a defining characteristic of language; systematicity is evident across languages, and iconic systematicity in particular is pervasive in ASL. This work makes clear that linguistic theorizing must consider iconicity and systematicity as a possible organizing principle of the lexicon and should be accounted for in psycholinguistic theories of how the lexicon is structured, used, and learned.

Materials and Methods

For the ASL analysis, we combined information about the phonological and iconic properties of 2,723 signs from the ASL-LEX database with recent data on semantic associations between signs. In the following section we briefly explain how these data were collected, but see Kezar et al. (2023), Caselli et al. (2017), and Sehyr et al. (2021) for more exhaustive details. To replicate these analyses in two spoken languages, English and Spanish, we drew on data from existing iconicity ratings, phonetic transcriptions, and semantic association datasets. For each of the more than 7 million pairs of signs in ASL-LEX, 26 million English pairs, and nearly 3 million Spanish pairs, we quantified iconicity, phonological associations, and semantic associations, using the procedures described below.

Summed Iconicity. For ASL, each of the 2,723 signs in ASL-LEX was evaluated by approximately 30 hearing non-signers, who determined how much the sign resembled its meaning. For example, they watched a video of the sign [CAT] and reported how much the sign resembled a cat on a scale of 1-7. These ratings were aggregated to determine the average iconicity rating for each sign. In the following analysis, for all pairwise combinations of signs in ASL-LEX, we added the average iconicity of the two signs in order to determine the summed iconicity of the pair. For English and Spanish, iconicity ratings were drawn from Winter, 2023 and Hinojosa, 2023, respectively. In these studies, native speakers of the languages (on average, $N_{English}$ = 10, $N_{Spanish}$ = 22.6 raters per word) were presented with words and asked to rate the extent to which words sounded like what they mean. As with the ASL data, the iconicity of word pairs was averaged across raters per word and then these averaged iconicity ratings were summed across word pairs. This variable was scaled for all analyses.

Phonological Associations. For ASL, trained, fluent-signing linguists annotated the phonological features of each sign in ASL-LEX. In cases where there were multiple sequential units (e.g., compounds, fingerspelled signs), we focus this analysis on the description of just the first constituent. Each possible pair of signs in the lexicon was then matched to determine how many phonological features the pair shared vs. diverged on, out of a maximum of 16. In some analyses, we binned phonological relatedness by labelling pairs that shared more than eight phonological features as "phonologically related." For the spoken languages, phonetic transcriptions of the words were drawn from LexOPS (English, Taylor, Beith, & Sereno, 2020) and EsPal (Spanish, Duchon et al., 2013). The Levenshtein edit distance in phonemes was then calculated for each word pair.

Semantic Associations. For ASL, we used semantic associations from a large-scale dataset of semantic associations between signs (Kezar et al., 2023). In this dataset, deaf ASL signers saw each of the 2,723 signs in the ASL-LEX database (Caselli et al., 2017; Sehyr et al., 2021), and were filmed producing the first three signs that came to mind. Up to 15 participants responded to each cue sign, yielding up to 45 associations per sign. The average age of first exposure to ASL was 2.44 years, (*Mdn* = 0*, SD* = 4.63, Range = 0 – 19). The complete ASL dataset included 113,883 semantic associations. For the following analyses, we removed responses that were impossibly short (<500ms, n = 909 responses excluded) or long (>6000ms, n = 714 responses excluded), included a corrupted video, or the response did not match a sign in ASL-LEX. For the statistical analysis, for every possible pair of signs in ASL-LEX, we determined whether the pairs were semantically related (i.e., at least one person freely associated the pair, n = 40,580) or not semantically related (n = 7,362,022).

For the two spoken languages, semantic associations were drawn from the Small World of Words (SWOW) project via their large-scale, crowd-sourced word association task. For these norms, participants were presented with over 13,000 cue words per language and asked to type the first three words that sprung to mind. Data collection procedures are described in more detail in De Deyne et al., 2018 (English) and Cabana et al., 2022 (Spanish). For each of the possible word pairs given the cue words, each word pair freely associated by at least *1* of the participants was considered semantically related. For English, this resulted in n = 292,514 semantically-related pairs and n = 26,527,752 semantically-unrelated pairs. For Spanish, the final dataset included n = 23,612 semantically-related pairs and n = 1,860,144 semantically-unrelated pairs.

Data Availability. The full list of word pairs along with their semantic associations (number of people who associated the pair) and their phonological associations (number of phonological features in common) as well as the code used in the statistical analyses are available at https://osf.io/5y6s4/.

Acknowledgments

Thanks to Lauren Berger for help with data collection, and Cindy O'Grady Farnady, Adele Daniels, Ruth Ferster, and Connor Baer for help with data collection and tagging. Thanks also to Amy Lieberman, Jennie Pyers, and the members of the Lava Lab and Lex Lab who helped us formulate these ideas over many lab meetings: Anna Lim, Amelia Becker, Tory Sampson, Michael Higgins, Paris Gappmayr, Jess Neri, Abigail Laughman, Lee Kazar, and Marshall Hurst.

References

1. Caselli, N. K., Sehyr, Z. S., Cohen-Goldberg, A. M., & Emmorey, K. (2017). ASL-LEX: A lexical database of American Sign Language. Behavior Research Methods, 49(2), 784–801. https://doi.org/10.3758/s13428-016-0742-0
2. Whitney, W. D. (1867). *Language and the Study of Language: Twelve Lectures on the Principles of Linguistic Science*. C. Scribner.
3. Saussure, F. de. (2011). *Course in General Linguistics* (P. Meisel, Ed.; W. B. E. by P. Meisel & H. Saussy, Trans.; p. 336 Pages). Columbia University Press.
4. Hockett, C. D. (1960). *The origin of speech*. W.H. Freeman.
5. Finegan, E. (2014). *Language: Its structure and use*. Cengage Learning.
6. Yule, G. (2022). *The Study of Language* (8th ed.). Cambridge University Press. https://doi.org/10.1017/9781009233446
7. Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. Proceedings of the National Academy of Sciences, 113(39), 10818-10823.
8. Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. Trends in Cognitive Sciences, 19(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013
9. Firth, J. R. (1957). A Synopsis of Linguistic Theory, 1930-1955.
10. Klima, E. S., & Bellugi, U. (1979). *The Signs of Language*. Harvard University Press.
11. Myklebust, H. R. (1960). *The psychology of deafness: Sensory deprivation, learning, and adjustment* (pp. xii, 393). Grune & Stratton.
12. Brentari, D. (2019). *Sign language phonology*. Cambridge University Press.
13. Börstell, C., Lepic, R., & Belsitzman, G. (2016). Articulatory plurality is a property of lexical plurals in sign language. *Lingvisticae Investigationes*, *39*(2), 391–407. https://doi.org/10.1075/li.39.2.10bor

14. Hwang, S.-O., Tomita, N., Morgan, H., Ergin, R., Ilkbasaran, D., Seegers, S., LEPIC, R., & Padden, C. (2016). Of The Body and The Hands: Patterned Iconicity for Semantic Categories. *Language and Cognition*, *9*, 1–30. https://doi.org/10.1017/langcog.2016.28
15. Östling, R., Börstell, C., & Courtaux, S. (2018). Visual Iconicity Across Sign Languages: Large-Scale Automated Video Analysis of Iconic Articulators and Locations. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00725
16. Padden, C., Meir, I., Hwang, S.-O., Lepic, R., Seegers, S., & Sampson, T. (2013). Patterned iconicity in sign language lexicons. *Gesture*, *13*. https://doi.org/10.1075/gest.13.3.03pad
17. Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a General Property of Language: Evidence from Spoken and Signed Languages. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00227
18. Occhino, C. (2017). An Introduction to Embodied Cognitive Phonology: Claw-5 Handshape Distribution in ASL and Libras. *Estudios Ingleses de La Universidad Complutense*, *25*, 69–104.
19. Köhler, W. (1929). *Gestalt psychology*. New York : H. Liveright. http://archive.org/details/gestaltpsycholog0000kohl
20. Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1651), 20130299. https://doi.org/10.1098/rstb.2013.0299
21. Martinez del Rio, A., Ferrara, C., Kim, S. J., Hakgüder, E., & Brentari, D. (2022). Identifying the Correlations Between the Semantics and the Phonology of American Sign Language and British Sign Language: A Vector Space Approach. Frontiers in Psychology, 13. https://www.frontiersin.org/articles/10.3389/fpsyg.2022.806471
22. Sehyr, Z. S., Caselli, N. K., Cohen-Goldberg, A. M., & Emmorey, K. (2021). The ASL-LEX 2.0 Project: A Database of Lexical and Phonological Properties for 2,723 Signs in American Sign Language. The Journal of Deaf Studies and Deaf Education, 26(2), 263–277. https://doi.org/10.1093/deafed/enaa038

23. Lepic, R., & Padden, C. (2017). *A-Morphous Iconicity*. https://doi.org/10.5281/ZENODO.495463
24. Jackendoff, R., & Audring, J. (2020). Relational Morphology: A Cousin of Construction Grammar. Frontiers in Psychology, 11. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.02241
25. Haslett, D. A., & Cai, Z. G. (2023). Systematic mappings of sound to meaning: A theoretical review. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02395-y
26. Bosworth, R. G., & 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*(6), 1573–1581. https://doi.org/10.1037/a0020934
27. Caselli, N. K., Emmorey, K., & Cohen-Goldberg, A. M. (2021). The signed mental lexicon: Effects of phonological neighborhood density, iconicity, and childhood language experience. Journal of Memory and Language, 121, 104282. https://doi.org/10.1016/j.jml.2021.104282
28. Ortega G., Morgan G. (2015c). The effect of iconicity in the mental lexicon of hearing non-signers and proficient signers: evidence of cross- modal priming. *Lang. Cogn. Neurosci.* 30 574–585. 10.1080/23273798.2014.959533
29. C Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407
30. 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*(12), 1443–1448. https://doi.org/10.1177/0956797612459763

31. Caselli, N. K., & Pyers, J. E. (2017). The Road to Language Learning Is Not Entirely Iconic: Iconicity, Neighborhood Density, and Frequency Facilitate Acquisition of Sign Language. *Psychological Science*, *28*(7), 979–987. https://doi.org/10.1177/0956797617700498
32. Gappmayr, P., Lieberman, A. M., Pyers, J., & Caselli, N. K. (2022). Do parents modify child-directed signing to emphasize iconicity? *Frontiers in Psychology*, *13*, 920729. https://doi.org/10.3389/fpsyg.2022.920729
33. Kezar, L., Pontecorvo, E., Daniels, A., Baer, C., Ferster, R., Berger, L., Thomason, J., Sehyr, Z. S., & Caselli, N. (2023). The Sem-Lex Benchmark: Modeling ASL Signs and Their Phonemes. *The 25th International ACM SIGACCESS Conference on Computers and Accessibility*, 1–10. https://doi.org/10.1145/3597638.3608408
34. Cabana, Á., Zugarramurdi, C., Lisboa, J. V., & Deyne, S. D. (2022). *The "Small World of Words" Free Association Norms for Rioplatense Spanish*. PsyArXiv. https://doi.org/10.31234/osf.io/w6hc8
35. De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7
36. Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1
37. Taylor, J. E., Beith, A., & Sereno, S. C. (2020). LexOPS: An R package and user interface for the controlled generation of word stimuli. *Behavior Research Methods*, *52*(6), 2372–2382. https://doi.org/10.3758/s13428-020-01389-1
38. Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones*. Language Linguistics Compass, 6*(10) 654-672. https://doi.org/10.1002/lnc3.361
39. Thompson, B., Perlman, M., Lupyan, G., Sehyr, Z. S., & Emmorey, K. (2020). A data-driven approach to the semantics of iconicity in American Sign Language and English. *Language and Cognition*, *12*(1), 182-202. https://doi.org/10.1017/langcog.2019.52
40. Winter, B., Woodin, G., & Perlman, M. (2023). Defining iconicity for the cognitive sciences. *OSF Preprints.* https://doi.org/10.31219/osf.io/5e3rc

Figures and Tables

Figure 2. Interaction effects between semantic relatedness and iconicity on the phonological distance for A) American Sign Language, B) English, and C) Spanish. Error bars show standard deviation around the mean.

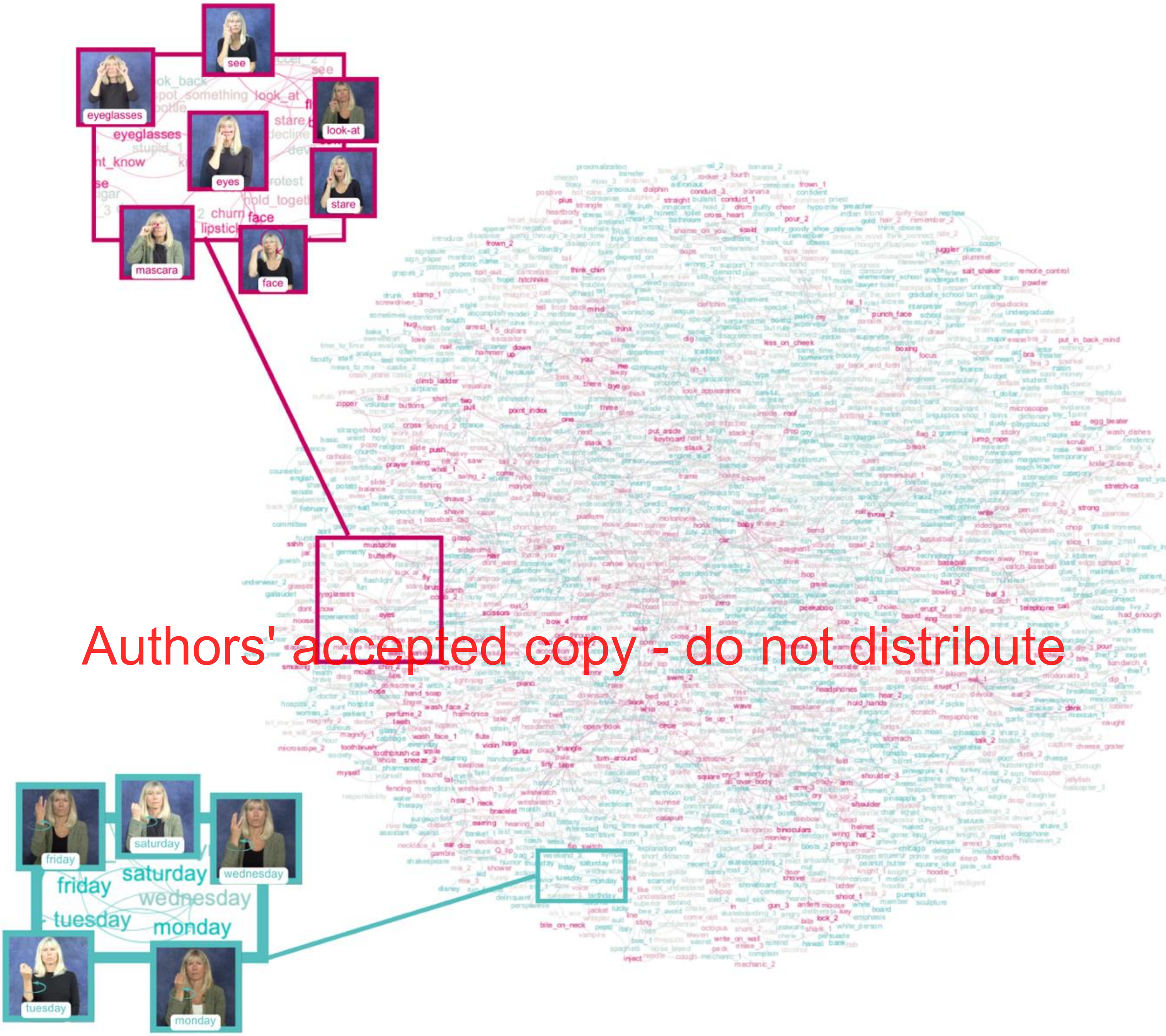

Figure 3. Systematicity in ASL: signs pairs that were semantically associated by at least two people, shared more than eight phonological features. Iconicity is encoded by color, with more iconic signs in magenta and less iconic signs in teal. Two clusters of interrelated signs are highlighted by cut-outs. Exemplar videos of all of these signs can be found at www.asl-lex.org.

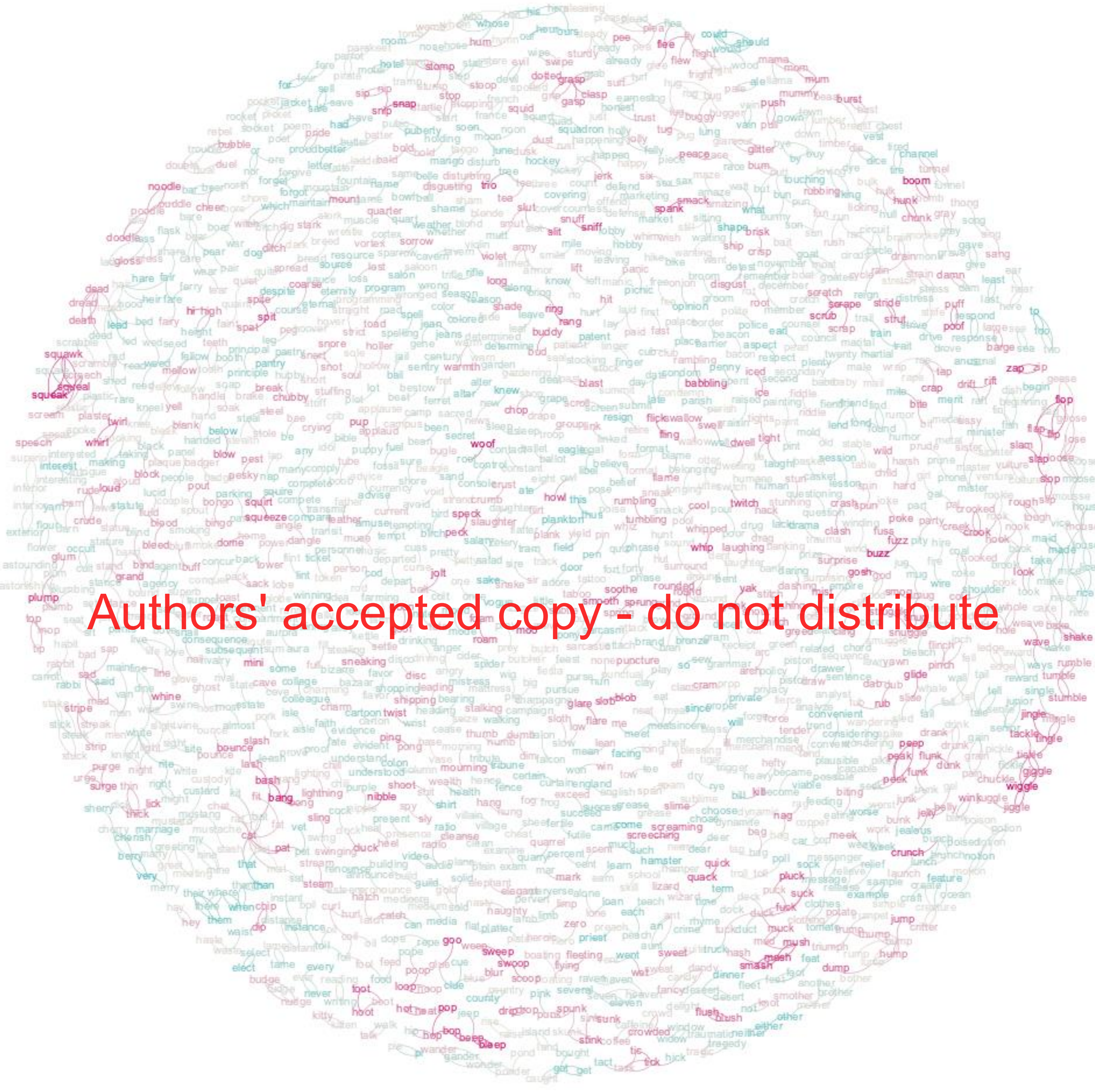

Figure 4. Systematicity in English. Word pairs are connected by an edge if they share at least half of their phonemes with one another and they were associated by at least 2 participants in SWOW. Iconicity is encoded by color, with more iconic words in magenta and less iconic words in teal.

Figure 5. Systematicity in Spanish. Word pairs are connected by an edge if they share at least half of their phonemes with one another and they were associated by at least 2 participants in SWOW. Iconicity is encoded by color, with more iconic words in magenta and less iconic words in teal.

Table 1. Effect of semantic relationships and summed iconicity on the number of differing phonological features in American Sign Language.

	Phonological Distance		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	9.83	9.80 – 9.87	**<0.001**
Semantic Relationship [Related]	-1.10	-1.14 – -1.06	**<0.001**
Summed Iconicity	-0.06	-0.09 – -0.03	**<0.001**
Semantic Relationships [Related] * Summed Iconicity	-0.13	-0.16 – -0.09	**<0.001**
Random Effects			
σ^2	4.97		
$\tau_{00\ Source}$	0.47		
$\tau_{00\ Target}$	0.47		
ICC	0.16		
N_{Source}	2723		
N_{Target}	2723		
Observations	7402602		
Marginal R^2 / Conditional R^2	0.001 / 0.160		

Table 2. Effect of semantic relationships and summed iconicity on the Levenshtein distance between phoneme strings in English.

	Phonological Distance		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.02	4.99 – 5.05	**<0.001**
Semantic Relationship [Related]	-0.16	-0.16 – -0.16	**<0.001**
Summed Iconicity	-0.22	-0.24 – -0.20	**<0.001**
Semantic Relationships [Related] * Summed Iconicity	-0.03	-0.04 – -0.03	**<0.001**
Random Effects			
σ^2	0.64		
$\tau_{00\ response}$	0.77		
$\tau_{00\ cue}$	0.67		
ICC	0.69		
N_{cue}	4829		
$N_{response}$	5554		
Observations	26820266		
Marginal R^2 / Conditional R^2	0.023 / 0.697		

Table 3. Effect of semantic relationships and summed iconicity on the Levenshtein distance between phoneme strings in Spanish.

	Phonological Distance		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.60	6.52 – 6.69	**<0.001**
Semantic Relationship [Related]	-0.32	-0.33 – -0.30	**<0.001**
Summed Iconicity	-0.12	-0.18 – -0.06	**<0.001**
Semantic Relationships [Related] * Summed Iconicity	0.01	-0.01 – 0.02	0.340
Random Effects			
σ^2	1.19		
$\tau_{00\ cue}$	1.23		
$\tau_{00\ response}$	1.23		
ICC	0.67		
N_{cue}	1374		
$N_{response}$	1372		
Observations	1883756		
Marginal R^2 / Conditional R^2	0.004 / 0.675		

Table 4. Network measurements of systematicity graphs in ASL, English, and Spanish. The Overall networks include all words that were systematically related (share at least half of their phonemes and were associated by at least 2 participants) to at least one other word. High-iconicity and low-iconicity subnetworks were determined by a median split on each language's iconicity ratings. Mean Degree refers to the mean number of edges each node in the network has. Mean Path Length is the mean number of steps in the shortest path between connected nodes in the giant component.

	ASL			English			Spanish		
	Overall	High Iconicity	Low Iconicity	Overall	High Iconicity	Low Iconicity	Overall	High Iconicity	Low Iconicity
% Words in Network	76%	55.1%	55.1%	38.6%	42.9%	35.9%	27.65%	28.0%	27.5%
Mean Degree	1.42	0.834	0.862	0.832	0.574	0.423	0.742	0.395	0.335
% Nodes in Giant Component	34.8%	17.4%	17.1%	1.93%	1.21%	1.13%	2.89%	2.56%	3.24%
Mean Path Length	16.457	3.57	4.51	5.19	3.45	3.06	1.61	1.22	1.197
Mean Clustering Coefficient	0.122	0.072	0.124	0	0	0	0.17	0	1.67