1

Iconicity in Expressives: An Empirical Investigation

MICHAEL GASSER, NITYA SETHURAMAN, AND STEPHEN HOCKEMA

1 Introduction

All comprehensive linguistic theories must confront the nature of the relationship between linguistic forms and the functions they perform. To what extent is this relationship constrained and where do the constraints come from? At the level of grammar, form-function relationships are apparently constrained in part by compositionality, and a range of linguistic or cognitive universals governing grammatical mappings have also been proposed. At the level of the lexicon, associations between form and function have, at least since Saussure (1966), been assumed to be mostly unconstrained. That is, the word form expressing a particular meaning in a particular language could be any phonotactically legal sequence of phonological segments and the convention for each such relationship must simply be memorized by learners of the language. It is in this sense that the form-function relationship in language is said to be mainly "arbitrary".

At the same time, it has often been recognized that within restricted regions of the lexicon, complete arbitrariness fails to hold. This is most obvious for onomatopoeia (e.g., *pitter patter*, *moo*), but it has also been attested for other categories of words in a wide range of spoken and signed languages (Hinton, Nichols & Ohala, 1994; Taub, 2001). In all of these cases, the forms of words are said to somehow "suggest" their meanings, or the

meanings are said to somehow "motivate" the word forms. The general term used for these phenomena is *iconicity*. Despite increased interest in iconicity in recent years, two fundamental questions remain unanswered:

1. What empirical means do we have for identifying a particular word as iconic or arbitrary?
2. How is it that the lexicon can be both arbitrary and iconic?

A possible answer to the second question is suggested by a completely different body of research, the study of lexical access and word recognition within psycholinguistics. Here it is recognized that the similarity of a word form to other forms plays a role in the speed at which the word is accessed during comprehension or production (Luce & Pisoni, 1998). Similarly, research on semantic speech errors during production (e.g., Fromkin, 1971) shows that words with similar meanings may interfere with one another. Since phonological similarity and semantic similarity both have an effect on language processing, it is worth considering whether the relationship between the two types of similarity also plays a role and whether communicative considerations could govern the sorts of form-meaning relationships that are favored or disfavored in languages.

In this paper, we provide a preliminary method of measuring the iconicity of words. This method relies on a new, formal definition of iconicity, one that distinguishes between what we call absolute iconicity, based on the direct relationship between form and meaning, and relative iconicity, based on the relationship between distance between forms and distance between meanings. We also propose an account of how arbitrariness and iconicity might be favored for different parts of the lexicon. We describe a preliminary study using Japanese and Tamil expressives, a lexical category reportedly characterized by a high degree of iconicity (Childs, 1994).

2   Expressives

Many languages of the world, especially in Africa, South Asia, Southeast Asia, and the Americas, have a lexical category referred to as *expressives*, also sometimes as *ideophones* or *mimetics* (Childs, 1994). While there are no necessary and sufficient conditions for defining this category cross-linguistically, a number of phonological, morphological, syntactic, semantic, and pragmatic properties characterize the expressive prototype. Expressive forms tend to be made up of a limited set of phonemes and to exhibit reduplication. Syntactically, expressives are often used in combination with a 'light' verb such as *say* or *do*. Expressives tend to denote adverbial meanings, and among the semantic categories often expressed by this category are movements (of the body or of objects); physical states; sounds and noises; speech patterns; sensations, emotions or mental states; personal ap-

pearance; facial expressions; and personality traits. In this paper, we focus on expressives in two unrelated languages, Japanese and Tamil, specifically on the extent to which they exhibit iconicity. Some examples of expressives in these languages are given in Table 1.

Table 1: Examples of Japanese and Tamil expressives

Japanese	Tamil
kirakira 'glitter, twinkle'	*labakk* 'suddenly'
giragira 'glare strongly'	*pacakk* 'suddenly'
girari 'glare momentarily'	*didikku* 'all of a sudden'

If it is true, as often claimed (e.g., Childs, 1994), that expressives are characterized by iconicity, why would this be? And do languages with expressives have other categories of words that are clearly distinguishable from expressives in terms of their arbitrariness? We return to these questions below. But first we need to be more precise about what we mean by "iconicity". What is it about words such as those in Table 1 that makes them iconic, and how would we demonstrate their iconicity in an objective manner?

3   Types of Iconicity

We consider words to be associations of what we will call **forms** and **meanings**. Both forms and meanings can be thought of as points in multi-dimensional spaces, though we may not have direct access to what the set of defining dimensions is. Given a meaning, a language user should be able to assign a form to it: this is production. Given a form, a language user should be able to assign a meaning to it: this is comprehension. Figure 1 illustrates these two processes for the simple case of two form and two meaning dimensions.
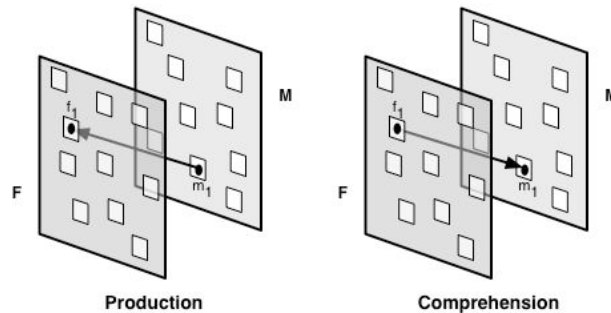

Figure 1. Production and Comprehension. Form space (F) and Meaning space (M) are each defined over only two dimensions. Each small square represents the form or meaning of a word. In production, the speaker begins with a meaning (m_1) and accesses a form (f_1). In comprehension, the addressee begins with a form (f_1) and accesses a meaning (m_1).

We define **absolute iconicity** to be the property of a set of words for which there is a similarity function that relates form and meaning (either in whole or part). In absolute iconicity, there is a correlation between one or more form dimensions and one or more meaning dimensions. The most obvious example is onomatopoeia, for which forms are intended as imitations of sounds in nature. Thus there appears to be a weak correlation between the vowel formants in conventional words for animal sounds (*moo*, *quack*, *cheep*) and the perceived formants in the sounds made by the animals. Sign languages offer many other instances of absolute iconicity, for example, the conventional use of the space in front of the signer for future times and the space behind the signer for past times. Note that the form-meaning function itself can vary from one language to another. Though we know of no such languages, it is at least conceivable that a sign language could assign the future to the space behind the signer and the past to the space in front of the signer. The point is that the assignment applies consistently across all of the words in the iconic set.

We define **relative iconicity** to be the property of a set of words for which there is a correlation between form similarity and meaning similarity. That is, rather than a similarity function relating meaning to form, relative iconicity is based on separate similarity functions relating forms to forms and meanings to meanings. For example, *glow, gleam*, *glint*, and *glimmer* seem to exhibit relative iconicity because their similarity of form (*gl-* in the salient initial position) corresponds to a similarity of meaning ('having to do with light of low intensity') and because there are relatively few other English verbs that begin with *gl-* and do not share this meaning. Note that with relative iconicity there need not be form-meaning similarity – this relationship could be completely arbitrary – as long as similar forms have similar meanings and similar meanings have similar forms. Note also that both kinds of iconicity are properties of groups of words, absolute iconicity because it is defined with respect to a similarity function over a set of form-meaning pairings and relative iconicity because it is defined with respect to the similarity between pairs within a group of words. In particular it makes no sense to say that a particular word is relatively iconic.

We define **arbitrariness** to be the absence of either absolute or relative iconicity. In this paper we will be mainly concerned with arbitrariness in the latter sense. For example, the words in the set *apple*, *pear*, *peach*, *apricot*, *plum*, and *cherry* appear to exhibit arbitrariness in the sense of absence of relative iconicity within the set since there is no obvious relationship between the similarity of forms and the *prima facie* similarity of associated meanings. For example, peaches and apricots are quite similar (relative to other objects or foods), but the forms of the words are as different as any

two pairs on this list. And the meanings of the only two nouns in the list that begin with vowels (in both cases followed by /p/) appear no more similar than any other pair.

A special kind of relative iconicity is what we call **anti-iconicity**, where expressions with similar meanings tend to have *dissimilar* forms. That is, there is a *negative* correlation between form similarity and meaning similarity. One of the goals of this paper is to demonstrate that anti-iconicity actually exists within some limited domains, but we illustrate here with a trivial example. Consider these common nouns for carnivores: *dog*, *wolf*, *fox*, *cat*. If we were to judge the similarities of the meanings of these words, relying either on their zoological classification or on their gross appearance and behavioral patterns[1], we might find dogs and wolves most similar and foxes more similar to dogs/wolves than they are to cats. But with respect to phonetic form, we find that *cat* and *dog* both consist of stop-vowel-stop sequences, whereas both *wolf* and *fox* consist of consonant-vowel-consonant-fricative sequences. That is, it would appear that within this set of words, form similarity correlates negatively with meaning similarity.

One problem with our informal examples of relative iconicity, arbitrariness, and anti-iconicity is that they have relied on our intuitive impressions of form and meaning similarity. For the definitions to be useful, we need explicit means of measuring form similarity and meaning similarity. With no constraints on what counts as similar, one could argue, for example, that in our fruit example, the presence of /p/ in words for fruits in the list that do not grow in clusters is evidence for iconicity within this list. We return to this issue in Section 6 below, offering a formal measure of form similarity and tentative approach to the trickier problem of meaning similarity.

In Figure 2, we abstractly illustrate absolute and relative iconicity and in Figure 3, anti-iconicity and arbitrariness.

---

[1] Note that these are not the only possible criteria for comparing the animals. On a scale of tameness, for example, the similarities would clearly turn out differently.
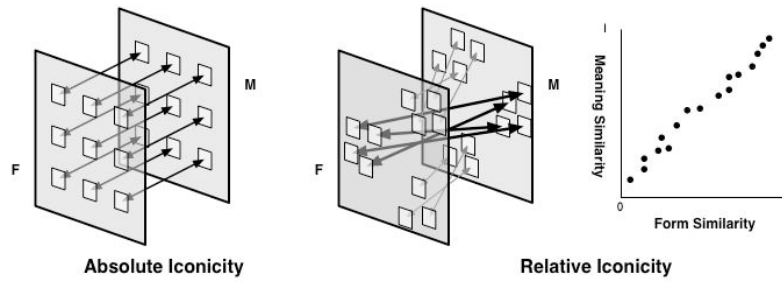
Figure 2. Absolute and Relative Iconicity. The arrows indicate associations between the forms and meanings of individual words. For the absolute iconicity example, there is a perfect correlation between both pairs of F and M dimensions. For the relative iconicity example, we highlight the connections for one cluster of words to show how similarity of form is associated with similarity of meaning. In the plot on the right, each point represents the form and meaning similarities for a single pair of words in the iconic domain.
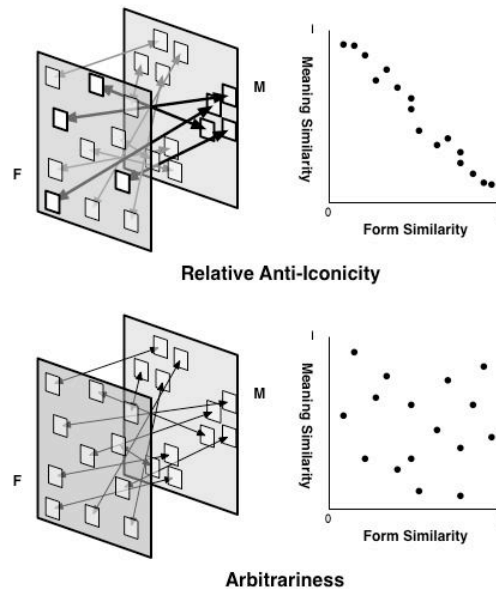

Figure 3. Relative Anti-Iconicity and Arbitrariness. For the anti-iconic example, one cluster of meanings is highlighted to show how its forms tend to be further apart than they would be by chance. As in Figure 2, in the plots on the right, each point represents the form and meaning similarities for a pair of words in the domain.

Given these definitions of relative iconicity and anti-iconicity, there are three possibilities for what is actually true of a particular natural language (different languages could differ in this regard):

1. The lexicon of the language could be arbitrary. Any apparent relative iconicity within a group of words would be spurious; if we randomly selected enough groups, we would eventually find what appeared to be iconicity in at least one of them.

2. The lexicon of the language could be relatively iconic. Any apparent arbitrariness within a group of words would be spurious; if we randomly selected enough groups, we would eventually find what appeared to be arbitrariness in at least one of them.

3. The lexicon could be arbitrary in some semantic or phonological regions and relatively iconic in others. This would require some definition of a region that is not circular; it would not be enough to show that a language exhibits iconicity for one set of semantically related words and arbitrariness for another set because one or the other could have happened by chance. That is, there would have to be some independent way of defining the regions where we find iconicity and those where we find arbitrariness.

We believe that this third possibility is likely, at least for certain languages, and in the rest of this paper we attempt to show that relative iconicity characterizes expressives and arbitrariness or anti-iconicity characterizes concrete nouns. But first we ask why this would be the case.

4 Advantages and Disadvantages of Iconicity and Arbitrariness

If we are right and the lexicons of natural languages exhibit both relative iconicity and arbitrariness, even anti-iconicity, why would this be? Are there areas within the lexicon that are more compatible with one or the other sort of relationship and, if so, is there a functional or cognitive explanation for this difference? In this section, we discuss several possible advantages and disadvantages of iconicity and arbitrariness for learning, production, and comprehension.

Iconicity would seem to have an advantage when it comes to the storage of forms in long-term memory and their retrieval during production. If the learner already knows one or more words in a domain that is known to be characterized by relative iconicity, then a new word in that domain should be easier to remember and to retrieve during production because of the similarity of its form to forms already learned. That is, if one knows that form f_1 has meaning m_1 and is then presented with a new form-meaning pair f_2-m_2, it should be easier to remember this pair knowing that the similarity

between f_1 and f_2 is reflected in the similarity between m_1 and m_2. It should also be easier to recall f_2 given meaning m_2 in production. For a possible advantage of arbitrariness in learning, see Gasser (2004).

For comprehension, we distinguish two different scenarios in which relative iconicity would be expected to have different consequences. Consider first a communicative situation in which the addresser (speaker or signer, S) has produced a word that the addressee (A) is unfamiliar with or has forgotten. If the word has phonetic properties that mark it as a potential member of the class of iconic words in the language, for example, if it features reduplication, then A may be able to narrow down the range of possible meanings for the unfamiliar form based on similar, familiar forms. For example, if A knows that *betabeta* means 'sticky' in Japanese, A may guess (correctly in this case) that the unfamiliar word *betobeto* has a similar meaning. In a context where the range of possible meanings for the unfamiliar word is constrained, this help may be all that is needed for A to figure out what S intended.

Now consider another type of communicative situation, one in which iconicity could interfere with comprehension. Say the context contains several potential referents all belonging to one relatively narrow semantic category, for example, gardening tools, fruits, articles of clothing, or internal organs. S refers to one member of the category with a noun that is familiar to A; for example, S says, "*Go out to the garage and get the **hoe**.*" Now assume that because of noise of one kind or another, the target word form is misheard. If the domain were characterized by relative iconicity, then the misheard form could be understand as a word for another member of the semantic category, for example, in the sentence above, for a shovel or a rake. Since the difference between a hoe and a shovel presumably matters in such a context, relative iconicity has the potential to *interfere* with communication. Arbitrariness or anti-iconicity, on the other hand, might have an advantage in such a situation. If the misheard form is a word in the language and its meaning is very different from that of the intended word, as would be expected with anti-iconicity, then there is little possibility of confusion with one of the other potential referents. In the example above, if *hoe* is misheard as *toe* or *hole*, then A is likely to realize that an error has occurred somewhere since these words don't refer to objects that one would retrieve from a garage. Rather than attempt to go and look for a toe or a hole, A would be expected to ask S for clarification. Compare this with what might happen if the English word for 'shovel' were *foe*.

Thus we see that both iconicity and arbitrariness or anti-iconicity could have advantages in communicative situations. Are there areas within the

lexicon where one of the other of these situations would be more likely to arise? We address this in the next section.

5 Where Should Iconicity and Arbitrariness Turn Up? Specific Predictions

We have argued that arbitrariness or anti-iconicity should be advantageous when an instance of a word has a set of potential referents that resemble one another but need to be distinguished. Thus we should expect arbitrariness or anti-iconicity in semantic domains where there are relatively rigid boundaries between the categories and the distinctions are crucial for communication. We believe that such semantic domains are most common within the category of basic concrete nouns: nouns referring to artifacts, to food, to geographical features.[2]

On the other hand, we should expect iconicity in semantic domains where the boundaries between the categories are not crucial for communication, where the disadvantage of possible confusion between similar meanings does not outweigh the advantage that iconicity should have for permitting the addressee to guess the meanings of unfamiliar words within the category. We believe that such semantic domains include those that are typical for expressives in languages that have them, domains such as manner of motion and bodily experience. For example, it may not be particularly important to keep distinct the meanings of groups of words such as *stumble/stagger/hobble* and *woozy/groggy/dizzy*.[3]

Our predictions then are:
1. Arbitrariness or even anti-iconicity should characterize particular semantic domains within the category of familiar concrete nouns.
2. Relative iconicity should characterize particular semantic domains within the category of expressives in languages that have them.

To test these predictions, we require a means of measuring the relative iconicity for a set of words and a means of isolating a set of concrete nouns and a set of expressives within candidate semantic categories. In the next section, we describe how we address each of these requirements.

---

[2] Although we don't discuss this notion further in this paper, labels also play a role in judgments of similarity (Sloutsky & Fisher, 2004), and anti-iconicity might cause the categories in such domains to be perceived as more distinct than they would otherwise.

[3] We make no claims about whether these English words are iconic by our definition of relative iconicity. We use them here only to illustrate the sorts of meanings that are conveyed by expressives in languages such as Japanese and Tamil.

6 Method

6.1 Iconicity Coefficient

We would like to know the relative iconicity for a set of words, that is, the correlation between the form similarity and meaning similarity for pairs of words in the category. The measurement of meaning similarity is the more difficult problem, and for the purposes of this paper, we simplify the problem by treating words as either inside or outside of particular semantic categories. This is equivalent to treating meaning similarity as a binary dimension, with pairs of words within the category in question having maximum similarity, pairs of words with one member inside and one outside the category having minimum similarity, and pairs of words outside the category having an undefined similarity.

In order to measure the relative iconicity of a category of words, we isolate a set of words within the category and a comparable set of words outside the category. We then calculate the mean form similarity for pairs of word forms within the category, s_w, and the mean form similarity between members of the category and words outside the category, s_b. Each of these quantities ranges from close to 0 (completely dissimilar) to 1 (identical). We define the **Iconicity Coefficient** (IC) to be s_w/s_b. If IC is significantly greater than 1.0 for a category, the category exhibits relative iconicity; if IC is significantly less than 1.0, the category is anti-iconic; and if IC $\approx$ 1.0, the category is arbitrary. The maximum IC value depends on the mean form similarity between randomly selected pairs of words. Given the algorithm for form comparison that is outlined below, this quantity is around .05. That is, as the form similarity of words within the semantic categories approaches 1, the IC for these very iconic sets approaches 20. Also note that, since absolute iconicity implies relative iconicity, IC is going to be maximal for an absolutely iconic language.

6.2 Form Comparison

The calculation of IC requires determining the form similarity of pairs of words. To accomplish this, we use an algorithm that attempts to align moras in the two words, yielding a score based on the phonetic similarity of the phones in corresponding moras and the degree to which the order of moras in the words is maintained in the alignment.

Each phone is represented as a vector of conventional phonetic features and phone similarity is 1.0 minus the normalized Euclidian distance between the phone vectors. The similarity between moras is the normalized sum of the similarities between corresponding phones in the moras. The

relative weight of vowels versus consonants within moras is controlled by a parameter that we currently set by hand at 0.7. This parameter should probably depend on the language. The mora alignment algorithm tries all possible alignments of moras in the two words, counting only the one with the highest score. The score for an alignment is the mean similarity of the aligned moras divided by the number of mora pairs in the alignment that appear in different orders in the two words. A mora in one word with no corresponding mora in the other word receives a similarity of 0.0. Table 2 shows some example form similarities for pairs of Japanese words and phone sequences (values range from 0.0 to 1.0).

Table 2: Examples of Japanese form similarity

betabeta / betobeto	.762		*betabeta / beta*	.125
betabeta / petapeta	.672		*betabeta / doron*	.050
betabeta / gitogito	.404		*betabeta / pin*	.016
betabeta / bettari	.222			

6.3 Selection of Words

For the expressives, we randomly selected 300 words from a dictionary of expressives for Tamil (Sataasivam, 1966) and an expressive thesaurus for Japanese (Chang, 1990). For the nouns, we relied on Japanese (Ogura & Watamaki, 1998) and Tamil (Sethuraman, ms.) versions of the MacArthur-Bates Communicative Development Inventory (Fenson et al., 1993; hereafter "CDI"), which lists the words most commonly known by infants and toddlers. We collected all of the words in the concrete noun categories, totaling about 350 items for each language.

For a baseline measure with which to compare results for particular categories, we calculated mean form similarity within the nouns and expressives and between the two categories for both languages (Table 3). Not surprisingly, in both languages, nouns and expressives are more similar to each other than to words in the other category. The results also show that Tamil forms are on average less similar to one another than Japanese forms are, probably due to the larger phonemic inventory for Tamil.

Table 3: Baseline form similarities

	Within nouns	Within expressives	Between nouns and expressives
Japanese	.0621	.0920	.0556
Tamil	.0427	.0372	.0280

7 Study 1: Iconicity and Anti-Iconicity in Broad Semantic Categories

We started by isolating relatively broad semantic categories. A thesaurus is organized in terms of a hierarchical ontology of categories, so, to the extent that we trust this ontology to accurately reflect how people group concepts, we can generalize results based on a thesaurus. For this study, we selected a subset of the most general categories from the Japanese expressive thesaurus (Chang, 1990). These included, for example, EMOTION and MOTION. For the Tamil expressives, we selected roughly comparable categories from the list of expressives in the Tamil expressive dictionary (Sataasivam, 1966). For the nouns, we selected several of the high-level categories that group the words in the CDI, for example, ARTICLES OF CLOTHING and FOOD ITEMS.

Our results show that the Iconicity Coefficient was not significantly different from 1.0 for any of the categories. That is, within broad categories grouping expressives and concrete nouns, we found no tendency towards iconicity or anti-iconicity. Note, however, that our hypotheses for the advantages of iconicity for expressives and anti-iconicity for concrete nouns both depend on the degree of semantic similarity within the categories we are examining. Perhaps the variation within broad categories such as EMOTION and FOOD ITEMS is too great for these advantages to be reflected. Given an unfamiliar expressive, there is presumably little advantage for the hearer to guess that it refers to some emotion on the basis of its similarity to other emotion words. The possibility of guessing that it refers to more specific emotions, such as anger or sadness, on the basis of its similarity to words in those categories, would be more useful. Similarly, there is little advantage for a misheard word for a food item to be less similar to *any* other food word than it is to other nouns. A form for a fruit that could not easily be confused with the names for other fruits, for example, would be more useful.

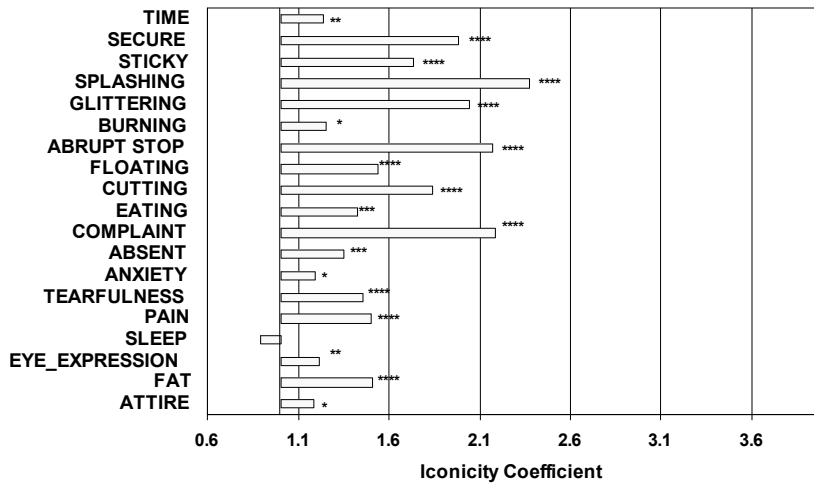In the next study, we therefore examined narrower semantic categories.

8 Study 2: Iconicity and Anti-Iconicity in Narrow Semantic Categories

In this study we examined the iconicity and anti-iconicity exhibited by the collected expressives and concrete nouns in meaning categories that are roughly at a level of specificity corresponding to the narrowest categories in a thesaurus, for example, words that describe pain of one sort or another. To collect such categories within the expressives, we selected 19 of the narrowest categories from the Japanese expressive thesaurus (Chang, 1990) that contained at least 8 words and 11 roughly-corresponding categories from the Tamil expressive list, using all common words in each category,

rather than only those from our original list of 300 expressives. For Japanese these categories included ATTIRE, SLEEP, PAIN, EATING, COMPLAINT, BURNING, SPLASHING, and STICKINESS; for Tamil they included ANGER, QUICKNESS, BABBLING, LAUGHTER, EATING, SOFTNESS, and SUDDENNESS. For the nouns, we informally selected categories roughly comparable in specificity. We avoided categories in which many of the words were recent borrowings or were morphologically complex because we felt these properties could affect the relative iconicity of the words (we return to these effects in the Conclusion section). Again, we made an attempt to include all "basic" members of these categories; there were between 8 and 13 of these in each case. The categories for both languages were FARM ANIMALS, INTERNAL ORGANS, NATIVE MUSICAL INSTRUMENTS, BODILY EXCRETIONS, SPICES, and WEATHER PHENOMENA.

To calculate the mean form similarity between words inside and outside a given category, s_b, we used all of the collected expressives and nouns in each case. For these two languages we take these two categories to be representative of the range of forms possible for mono-morphemic words (verbs in both languages are always inflected). For each category, we wanted to know whether its Iconicity Coefficient (IC) was significantly greater or less than 1.0, whether the within-category form similarity differed from the similarity of those words to the range of all words included in the sample. We used a Monte Carlo simulation to assess the significance in each case. This technique works by randomly picking words to place in hypothetical meaning categories within the larger lexical categories (nouns or expressives) and then using the distribution of the ICs of these categories to determine how likely the ICs we actually obtained would have occurred by chance.

Japanese Expressives

Label	
TIME	**
SECURE	****
STICKY	****
SPLASHING	****
GLITTERING	****
BURNING	*
ABRUPT STOP	****
FLOATING	****
CUTTING	****
EATING	***
COMPLAINT	****
ABSENT	***
ANXIETY	*
TEARFULNESS	****
PAIN	****
SLEEP	
EYE_EXPRESSION	**
FAT	****
ATTIRE	*

0.6 1.1 1.6 2.1 2.6 3.1 3.6

Iconicity Coefficient

Tamil Expressives

Label	
ORNAMENT_NOISE	****
UNEVEN	*
SUDDEN	****
SOFT	*
HEALTHY	
EAT	**
HEARTBEAT	**
LAUGH	
BABBLE	****
QUICK	
ANGRY	****

0.6 1.1 1.6 2.1 2.6 3.1 3.6
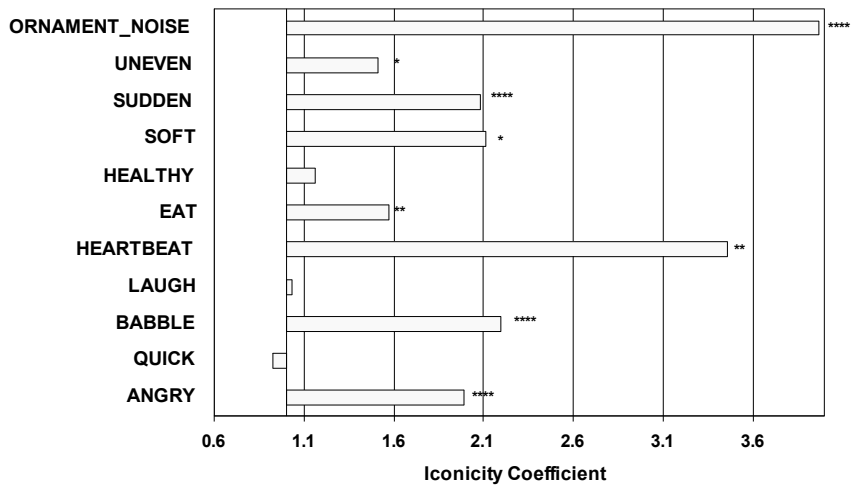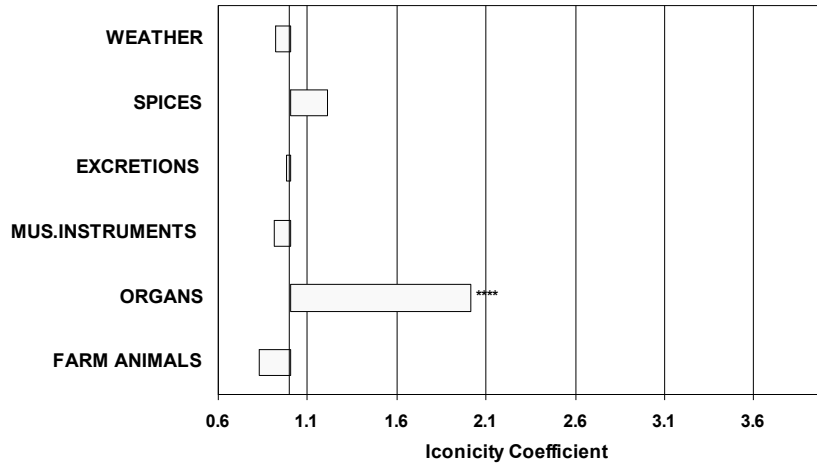
Iconicity Coefficient

Figure 4. Iconicity Coefficients for Japanese and Tamil Expressives
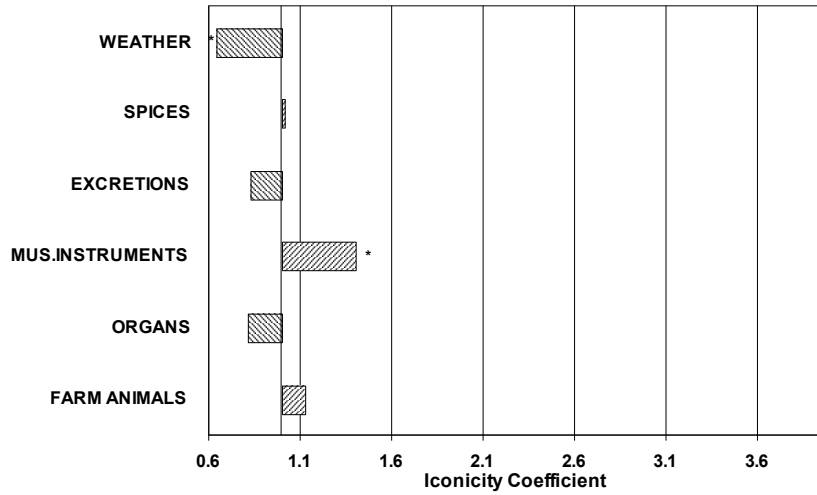
Japanese Nouns

Tamil Nouns

Figure 5. Iconicity Coefficients for Japanese and Tamil Nouns and Expressives

The results for expressives and nouns in each language are shown in Figures 4 and 5. In each figure, the x-axis indicates the IC for each of the categories of words represented by the horizontal bars. Recall that arbitrariness corresponds to an IC not significantly different from 1.0, relative iconicity to an IC significantly greater than 1.0, and anti-iconicity to an IC significantly less than 1.0. In the figures, categories with ICs significantly greater or less than 1.0 are indicated with asterisks next to the bars.[4]

For the expressives, we see that 8 of the 11 Tamil categories and 18 of the 19 Japanese categories are iconic. For the nouns, although half of the categories have ICs of less than one, only 1 of the 6 Tamil categories is significantly anti-iconic and none of the Japanese categories is. On the other hand, one Tamil and one Japanese category are significantly iconic. Thus for the expressives we have clear evidence, albeit from a relatively small sampling of categories, that words within narrow semantic categories exhibit relative iconicity and for concrete nouns we have some evidence, albeit from an even smaller and less rigorously selected sample, that words within narrow semantic categories tend not to exhibit relative iconicity.

There are two rather striking exceptions to this latter conclusion among the sample, one for Tamil and one for Japanese. For Japanese, the relatively high form similarity within the category of internal organ nouns is clearly related to the presence of the syllable *zoo* at the end of three of the words. All three of these are early borrowings from Chinese (as is about one-third of the Japanese lexicon), so this syllable probably behaves like a morpheme for Japanese speakers (although the syllable that remains if *zoo* is deleted probably does not). Thus, this set of words may not have satisfied our criterion of mono-morphemicity. On the other hand, we have no explanation for the high form similarity within the category of Tamil musical instrument words.

9   Discussion and Conclusion

The main contribution of our study is the demonstration that it is possible to quantify and measure relative iconicity, given agreed-on measures of phonetic and semantic similarity. Our more specific conclusions must be seen as somewhat tentative because of the small sample of words from each lexical category and questions about how we performed meaning comparison. These conclusions include the following:

---

[4] Asterisks indicate significance at the .05 (*), .01 (**), .001 (***) and .0001 (****) levels. If a Bonferroni correction is made, the significance cutoff rises: bars marked with a single asterisk will no longer be significant.

1. Expressives in Japanese and Tamil exhibit more relative iconicity than concrete nouns.
2. As the size of an expressive semantic category gets smaller, the likelihood of relative iconicity increases. If this effect proves to hold for other languages, the communicative advantages that we propose for iconicity may only apply for very narrow categories.
3. In very narrow semantic categories within concrete nouns, anti-iconicity is also possible. If this result generalizes to other languages, the communicative advantages that we propose for arbitrariness or anti-iconicity may only apply for narrow categories.

In future work we will need to address both form and meaning comparison. Our form comparison algorithm may be missing aspects of perceived similarity because it fails to take into account reduplication and we also have only guessed at the appropriate settings for the small number of parameters that govern the similarity measure, for example, at the relative weight to assign to the onset and the nucleus within a mora and at the relative weight to assign to phonetic-feature differences within segments. What is needed is experimental validation of the measure we are using.

Meaning comparison is clearly much more problematic. In order to select semantic categories and to select the words in the categories, we relied on dictionaries and thesauri, but these sources rely in turn on the intuitions of their authors. We are currently exploring ways of using subject similarity ratings as an alternative.

We also will have to address several larger theoretical issues. One of these concerns the whole notion of "meaning similarity". On any account, meaning is multi-dimensional and a person's judgment of the similarity between two words may be based on different dimensions or combinations of dimensions, for example, whether one compares animals by their overall SHAPE or their TAMENESS. Crucially, which dimensions come into play may depend on the communicative context. Our preliminary investigations suggest that certain dimensions may favor iconicity whereas others favor arbitrariness/anti-iconicity. For nouns, a dimension such as POLITENESS or REGISTER may be more constraining for words that are maximally *dissimilar* on that dimension but similar on all others (for example, *feces* and *poop*): such words should be dissimilar in form. In future work, we will need to examine how meaning similarity depends on context and whether certain specific communicative contexts are more likely to lead to the effects we predict for expressives and concrete nouns.

Lastly, one question we have barely touched upon is how composition interacts with arbitrariness/anti-iconicity, e.g., in sets such as *blackberry, strawberry, blueberry*. Here we find form similarity (a common morpheme) for semantically related words, a relationship that runs counter to what we would predict for concrete nouns (and leads to our anomalous finding for

Japanese nouns in the INTERNAL ORGAN category). It may be that in such cases, addressees learn to use the morpheme overlap to help them recognize the gross meaning category and to use the *dissimilarity* of the non-overlapping portions to avoid confusion among subcategories within this gross category. We plan to investigate this possibility using native-speaker judgments of form similarity for mono-morphemic and poly-morphemic pairs of words.

References

Chang, A. C. (1990). *Gitaigo-giseigo bunrui youhou jiten. A thesaurus of Japanese mimesis and onomatopoeia: Usage by categories*. Tokyo: Daishuukan Shoten.

Childs, G. T. (1994). African ideophones. In Hinton, Nichols, & Ohala (Eds.), *Sound symbolism*, pp. 179-204. Cambridge: Cambridge University Press.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego: Singular Publishing Group.

Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*, 27-52.

Gasser, M. (2004). The origins of arbitrariness in language. *Proceedings of the Annual Conference of the Cognitive Science Society*, *26*.

Goldstone, R.L., Medin, D.L., & Halberstadt, J. (1997). Similarity in context. *Memory & Cognition*, 25 (2), 237-255.

Hinton, L., Nichols, J., & Ohala, J. (Eds.) (1994). *Sound symbolism*. Cambridge: Cambridge University Press.

Luce, P.A. & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, *19*, 1-36.

Ogura, T. & Watamaki, T. (1998). Research and development of early language development inventory. Research project, Grant-in-aid for Scientific Research, Kobe University (in Japanese).

Peirce, C. S. (1998). *The essential Peirce: Selected philosophical writings: Volume 2*. Bloomington, IN: Indiana University Press.

Sataasivam, M. (1966). *Olikkurippakaraati* (Dictionary of onomatopoeic expressions). Madras: Paari Nilayam.

Saussure, F. de (1966). *Course in general linguistics* (tr. W. Baskin). New York: McGraw-Hill.

Sethuraman, N. (ms.) Early language inventory: Tamil. Manuscript.

Sloutsky, V.M., & Fisher, A.V. (2004). Induction and categorization in young children: a similarity-based model. *Journal of Experimental Psychology: General*, *133*, 166-188.

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.