

Predicting iconicity — coding scheme v01

Definitions: basic forms of lexical iconicity (based on [1–3])

Iconicity. The resemblance-based mapping between aspects of form and aspects of meaning. Such mappings can be seen as perceptual or motor analogies linking form and meaning.

Imagic iconicity. Imagic iconicity involves within-modality perceptual analogies, as when a spoken word mimics a sound in the world. Form and meaning can be said to share substance.

Diagrammatic iconicity. Diagrammatic iconicity involves cross-modal perceptual analogies, as when repeated syllables express the idea of repetition, or a vowel contrast maps onto a size contrast. Form and meaning can be said to share structure.

(Within diagrammatic iconicity, we can distinguish two subtypes. *Gestalt iconicity* is when word structure maps onto event structure: for instance, when a closed monosyllable refers to a single event with aspectual closure. *Relative iconicity* is when related forms map onto related meanings: for instance, when in a set of words, a vowel contrast like /i:o/ serves as a perceptual analogy for a size contrast /small:big/.)

Overall goal

Words differ in the degree to which they present transparent associations between form and meaning. Even within word classes traditionally seen as iconic (e.g., ideophones), there are large differences in ‘guessability’, with a word like Japanese *hisohiso* ‘whispering’ guessed correctly often but some other ideophones performing at chance level or even below. Prior work has been content to conclude that, on average, people can guess learn the meaning of ideophones at levels (just) above chance [4–6]. Here, the goal is to get a grip on the internal variance in such tasks. Which ideophones are easier to guess and why?

The hypothesis to be pursued here is that ideophones that are more transparently iconic are easier to guess and learn. The key question is how to operationalise the notion of iconicity. Here, we do this by focussing on a number of fairly general form-meaning associations independently attested in the literature. We code ideophones for aspects of their form (F) and meaning (M). When form is congruent with meaning (for instance, when a reduplicated ideophone form is paired with a meaning that involves repetition), this means that aspect of the form-meaning association is iconic. Because form-meaning associations are largely independent, they yield several separate measures of perceptuomotor analogies, and they can be combined into a cumulative measure of form-meaning congruency. Both the separate features and the cumulative measure can be inspected for the degree to which they predict the experimental guessability of ideophones (as established in prior studies).

Coding scheme: design considerations

Some of the form and meaning characteristics can be fairly straightforwardly coded in binary terms (e.g., reduplication: present or absent). Others may make more sense as continuous scales (e.g., magnitude: from small to large). For continuous measures, however, it is harder to check raw coding consistency across coders, and they might have to be dichotomised again later. If binary features are preferred, it should be kept in mind they may represent an oversimplification.

Form characteristics

F_redup	does the form feature full or partial reduplication? (1/0)
F_monosyllabic	is the form a monosyllable? (1/0)
F_partialredup	does the form feature partial reduplication? (1/0)
F_closedsyllable	does the form end in a closed syllable? (1/0)
F_vowelquality	are the vowels in the form mostly light (0), even (1), or mostly dark (2)? <i>Light vowels are /e ε i/, and dark vowels are /o ɔ u/. This includes the vowels' nasal variants.</i> <i>If an item has, for example, one light and one dark vowel, it is coded as even (1). If an item has one light vowel and two dark vowels, it is coded as mostly dark (2).</i> <i>Vowels /a ʌ ə i/ are considered neutral and do not influence the score: if an item has one light vowel and one neutral vowel, it is still coded as light (0).</i>
F_voice	are the consonants in the form mostly voiceless (0), even (1), or mostly voiced (2)? <i>Here, voiceless consonants are the voiceless plosives and fricatives. Voiced consonants are the voiced plosives, fricatives as well as nasal consonants. Other (semi-)consonants, such as laterals or glides, are considered neutral and do not influence the score. The score is built up the same way as in F_vowelquality.</i>
F_finalvowel	does the form end in a long final vowel? (1/0)
F_intonation	is the form spoken with relatively low (0), neutral (1), or relatively high (2) intonation compared to other items spoken by the same person?

Meaning measures

M_sound	is the meaning in the domain of sound? (1/0)
M_distribution	is there a sense of repetition, iterativity and/or distribution in space or time? (1/0)
M_irregular	is there a sense of irregularity, unevenness, chaos?
M_weight	where is the meaning on the following scale: small/light/hard (0), neutral (1), big/heavy/dark/soft (2)?
M_punctual	is there punctual (one-off) aspect? (1/0)
M_durative	is there durative (continuous, ongoing) aspect? (1/0)
M_abrupt	does the meaning involve an abrupt ending? (1/0)

Derived measures

FM_congruent_modality	M_sound = 1
FM_congruent_rep	F_redup = 1 && M_distribution = 1
FM_congruent_punct	F_monosyllabic = 1 && M_punctual = 1
FM_congruent_irregular	F_partialredup = 1 && M_irregular = 1
FM_congruent_closure	F_closedsyllable = 1 && M_abrupt = 1
FM_congruent_durative	F_finalvowel = 1 && M_durative = 1

FM_congruent_weight	F_vowelquality = 0 && M_weight = 0 OR F_vowelquality = 2 && M_weight = 2
FM_congruent_magnitude	F_intonation = 0 && M_weight = 2 OR F_intonation = 2 && M_weight = 0
FM_congruent_voice	F_voice = 0 && M_weight = 0 OR F_voice = 2 && M_weight = 2