When semantics aids phonology: a processing advantage for iconic word forms in aphasia

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Title:

When semantics aids phonology: A processing advantage for iconic word forms in aphasia.

Authors:

Lotte Meteyard[1]

Emily Stoppard[1]

Dee Snudden[1]

Stefano F. Cappa[2]

Gabriella Vigliocco[3]

*corresponding author

Affiliations

1: Department of Clinical Language Sciences, University of Reading, UK.

Tel: +44(0)118378 8142   Email: l.meteyard@reading.ac.uk

2:IUSS Pavia and Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy

3: Research Department of Experimental Psychology, University College London, UK

Abstract

Iconicity is the non-arbitrary relation between properties of a phonological form and semantic content (e.g. "moo", "splash"). It is a common feature of both spoken and signed languages, and recent evidence shows that iconic forms confer an advantage during word learning. We explored whether iconic forms conferred a processing advantage for 13 individuals with aphasia following left-hemisphere stroke. Iconic and control words were compared in four different tasks: repetition, reading aloud, auditory lexical decision and visual lexical decision. An advantage for iconic words was seen for some individuals in all tasks, with consistent group effects emerging in reading aloud and auditory lexical decision. Both these tasks rely on mapping between semantics and phonology. We conclude that iconicity aids spoken word processing for individuals with aphasia. This advantage may be due to a stronger connection between semantic information and phonological forms.

1.0 Introduction

The purpose of communication is to convey meaningful messages. Human language systems achieve this by associations between linguistic forms (spoken or signed words) and meanings (aspects of experience) (Ramscar et al., 2010). The mapping from linguistic forms to meaning during comprehension, and from intended meaning to linguistic forms in production, is carried out effortlessly and very efficiently. This is despite the presence of arbitrariness, the fact that the linguistic form does not provide *direct* cues to meaning in such a mapping.

Arbitrariness of the mapping between form and meaning has been long argued to be a foundational feature of human language systems (Saussure, 1916; Miller & Johnson-Laird, 1976). There is nothing inherent in the sound form "cat" that cues the meaning *cat*: a fluffy, four-legged predatory household pet with whiskers, night vision and a long tail. Form-meaning mappings arise from convention and, except for historical precedent, it could just as easily have been the sound "dog" that cues the meaning *cat*. Arbitrariness in the mapping has been argued to be one key aspect of the referential problem in word learning (Ramscar et al., 2010) – how does a child learn linguistic symbols, mapping objects and events in their environment to an arbitrary word form? Neurobiological models of language uphold the separation between form and meaning, with phonological and conceptual/semantic systems supported by largely separate brain networks (Price, 2012; Binder, Desai, Graves & Conant, 2009). For adults, word finding difficulties are one of the most ubiquitous complaints both of ageing adults (Burke & Shafto, 2008) and those with any form of acquired damage to language networks (Shewan & Kertesz, 1980). One reason for this may be that during production a unique phonological form has to be retrieved (e.g. Levelt, 1992); the arbitrary connection from semantics to phonology may be one reason why word

retrieval is so sensitive to minimal changes in the efficiency of language processing.

However, form-meanings mappings are not always arbitrary. Iconic relationships between form and meaning are widespread in both spoken and signed languages (Perniss, Thompson & Vigliocco, 2010; Schmidtke, Conrad & Jacobs, 2014). Iconicity refers to there being a non-arbitrary resemblance between the signifier (the word) and what is being signified (the concept) (Fischer & Nänny, 1999). For signed languages, iconicity is ubiquitous at lexical and sentential levels (Taub, 2001). At lexical levels, it describes the presence of an imagistic relationship between some manual and non-manual properties of the form (mouth, face and signer's body) and visual and motoric characteristics of what is being signed (Perniss, Thompson & Vigliocco, 2010). For example, the British Sign Language (BSL) sign BELT incorporates the action of putting a belt around the waist using a "C" handshape with both hands (Thompson, Vinson & Vigliocco, 2010). Iconicity in lexical signs undergoes conventionalization; arbitrary signs (e.g., MOUSE, produced with a curved index finger rotating on the signer's nose) are present along with more iconic signs. Goldin-Meadow (2005) presents an analysis of homesign systems, gestural communication used by deaf children not exposed to conventional spoken or signed languages. These children produce a broad variety of iconic signs, presumably because their communication needs to be transparent to their parents (Goldin-Meadow, 2005). However, they do not use all possible manual variations (e.g. handshapes) that are available. Instead, they set up consistent form-meaning pairings – i.e. morphemes – which means the iconic signs are not always precisely mapped onto their meaning (e.g. using a fist to denote handling a balloon string and an umbrella handle, which would not be identically held in real-life; Goldin-Meadow, 2005). This illustrates a movement from iconic to arbitrary signs. Such

"conventionalisation" in the home-sign systems of individual children may mirror processes seen in language evolution more broadly (Botha, 2007).

Among spoken languages, the degree of iconicity in the phonological form of words differs greatly. Outside the Indo-European language family, we find that iconic mappings are well represented in virtually all sub-Saharan African languages, some of the Australian Aboriginal languages, Japanese, Korean, Southeast Asian languages, indigenous languages of South America, and Balto-Finnic languages (Perniss et al., 2010). In these languages, in addition to more direct acoustic links, these iconic, sound-symbolic mappings evoke sensory, motor, or affective experiences or characterize aspects of the spatio-temporal unfolding of an event. In fact, the majority of sound-symbolic words refer to events or states in which sound is not essential. That is, properties of experiences – including visual, tactile, as well as mental and emotional experiences – may systematically correspond to properties of vowels and consonants, and their patterns of combination (e.g. reduplication) (Hamano, 1998).

Indo-European spoken languages, such as English, tend to have a lesser degree of iconicity in phonological forms. Here, iconicity tends to be for acoustic experiences (e.g. "splash", "trill", "croak", "pitter-patter", "bling", "moo", "whoof"; Perniss, Thompson & Vigliocco, 2010) although it is not limited to that domain. Iconic mappings can also arise when there are consistent *relationships* between particular forms and particular meanings; these statistical regularities then provide a consistent mapping between form and meaning (Monaghan, Shillcock, Christiansen, & Kirby, 2014). These regularities can be seen in phonoaesthemes, in which typically word initial or word final consonant clusters correlate with a certain meaning, e.g. "glitter", "gleam" and "glow" having a "*gl*" onset that maps to the meaning 'low intensity light' (Firth, 1930; Perniss, Thompson & Vigliocco, 2010). For more

abstract dimensions, mappings between vowel quality and the concept of size are present across spoken languages, for example, with high vowels ("ee") being associated to small sizes and low vowels ("aa") associated to large sizes (Hinton, Nichols, Ohala, 1994). It has been argued that abstract words cannot make use of sound-symbolism because they are not tied to specific aspects of experience, but other aspects of phonology may provide cues to their meaning (Reilly & Kean, 2006; Reilly et al, 2012). High and low imageability words differ on a number of phonological variables, with abstract words being longer, more derivationally complex and having fewer phonological neighbours (Reilly & Kean, 2006; Westbury & Moroschan, 2009). When asked to make a semantic judgement (concrete/abstract) about nonword stimuli, individuals consistently rate longer and inflected words as more abstract (Reilly et al., 2012). This sensitivity is preserved in semantic dementia with patients misclassifying concrete 3 syllable words (e.g. professor) more often than 1syllable words (e.g. bake) when making a forced choice concrete/abstract decision (Reilly et al., 2007). In this context, iconicity can be seen as part of the systematic way that linguistic forms cue aspects of meaning (Ramscar et al., 2010).

In English there is overall more systematicity than expected by chance. Monaghan, Shillcock, Christiansen & Kirby (2014) carried out a corpus analysis in English to explore how systematic the relationship is between form and meaning. They correlated measures of phonological and semantic similarity across words. As well as finding that English was more systematic than predicted by purely arbitrary mappings, words acquired earlier (earlier age of acquisition) were more systematic than words acquired later, suggesting a benefit of iconic mappings during language acquisition. A similar finding is present for BSL in parental report data (Thompson, Vinson, Woll & Vigliocco, 2012), with iconic signs being acquired earlier than less

iconic signs by children aged 11-24 months. In an analysis of the spoken output of a child learning German, onomatopoeic words (i.e., 'lala' for music or 'bow-wow' for dog) were shown to bootstrap vocabulary growth. A sharp increase in produced onomatopoeic words between 0,8 and 0,11 months preceded a more general vocabulary spurt from the age of 1 year 1 month (Laing, 2014).

Further empirical evidence for an iconic advantage for language learning comes from experiments with both children and adults. English 3 year olds are better able to learn Japanese verbs with iconic (sound-symbolism) properties (Kantartzis, Imai & Kita, 2011), and Japanese 3 years olds are better able to generalize the meaning of novel verbs if they have iconic properties (Imai, Kita, Nagumo & Okada, 2008). In adults, Nygaard, Cook and Namy (2009) used a vocabulary task for native speakers of American English to learn Japanese words. Learners were sensitive to consistency in form-meaning mappings in a language with which they had no prior experience. When Japanese words were paired to their correct meanings in English (rather than randomly paired), they were responded to more quickly and more accurately over learning blocks. Similar benefits of iconic mappings were shown by Kovic, Plunkett & Westermann (2010) in a combined behavioural and EEG study, in which participants had to map two nonsense words to novel object pictures. They found that participants were faster to respond when the mapping was iconically congruent (e.g. 'mot' mapped to a curvy object, 'riff' matched to a pointy object). The most robust ERP response was an increased negative wave between 140-180ms for congruent conditions at occipital sites, interpreted as reflecting early processes of auditory-visual integration.

Regarding processing, Thompson, Vinson and Vigliocco (2009, 2010) have shown that iconicity in sign forms affects sign recognition, suggesting that whenever

there is iconicity in the sign, signers cannot avoid retrieving aspects of the semantics, regardless of task. In spoken English, Westbury (2005) presented CVC words and nonwords in a spiky or curvy frame for lexical decision. Words were made up of continuants (e.g. "mime"), plosives (e.g. "cope) or a mixture of the two (e.g. "food); mirroring this, nonwords were also made up of continuants (e.g. "nool"), plosives (e.g. "dibe") or a mixture of the two (e.g. "nool"). Reaction times for lexical decision showed an interference effect for nonwords, with longer latencies when the frame was incongruent with the word structure (e.g. curvy frame with a plosive item); this effect was replicated when subjects made a decision on a single letter, rather than whole word/nonword items. Connell and Lynott (2014) explored the salience of a particular sensory modality and how that may cue meaning. They found that strongly visual words (e.g. cloudy) were responded to more quickly than weakly visual words (e.g. salty) during visual lexical decision and reading aloud, which both direct attention towards vision. Similarly, strongly auditory words (e.g. noisy) were responded to more quickly than weakly auditory words (e.g. salty) during reading aloud, as this task also directs attention towards auditory information. These effects indicate that cues to meaning are routinely processed and there may be automatic detection and *use* of form-meaning or modality-meaning consistencies whenever they are present.

The data from both spoken and signed languages indicate that iconicity confers some benefit for word learning and word processing, but how do these effects come about? One possibility is that for iconic words there are additional links between phonological and semantic networks. That is, iconicity provides an opportunity for greater embodiment in language processing (Meteyard, Cuadrado, Bahrami & Vigliocco, 2012), with the word form directly triggering aspects of modality specific experience. Brain imaging results provide some support for this

hypothesis. In a study using semantic categorization as task with Japanese participants, Hashimoto et al. (2006) presented nouns (animal and bird names), animal and bird calls, onomatopoeic sounds (imitations of animal and bird calls) and control sounds (white noise and tones). Results showed greater activation for onomatopoeic sounds than for nouns and environmental sounds in the left and right superior temporal sulcus (STS) and the right inferior frontal gyrus. The involvement of the right STS for iconic items was replicated by Kanero et al. (2014), who asked participants to judge the match between visually presented Japanese words and video clips. Sound-symbolic words were compared to non-sound symbolic verbs and adverbs (Experiment 1). When sound-symbolic words were matched to the videos, they found greater activation in the right posterior STS. The authors argued that the right STS activation evidenced iconic word forms as being processed both as lexical items and as non-lexical, sound symbols. Revill, Namy, DeFife and Nygaard (2014) compared sound symbolic and non-sound symbolic words (from a set of typologically diverse languages, unknown to participants) during a meaning selection task based on form/motion antonyms. Sound symbolic words were associated with higher levels of activation in the left intraparietal sulcus. Overall, these data point to the possible role of areas involved in cross-modal integration beyond those engaged in semantic processing (Binder et al., 2009). While the retrieval, combination and/or integration of information from different sources is a general aspect of semantic processing (Patterson, Nestor & Rogers, 2007), iconic words may benefit from more direct mappings to sensory domains (e.g. phonological to acoustic, phonological to visual, orthographic to visual) than non-iconic words. That is, iconic words engage additional processing networks that are 'outside' canonical language networks (e.g. as evidenced by greater right hemisphere activation for these kinds of words).

These results entail a prediction about the possible effects of brain lesions, i.e., we might expect iconic words to be, in general, less vulnerable than non-iconic words when left-hemisphere neurological networks underscoring language processing are damaged. Thus, individuals who have aphasia following stroke may show more accurate performance on production and comprehension of these words. Two previous studies have examined iconicity effects in deaf signers of British Sign Language (BSL) who have suffered strokes (Marshall, Atkinson, Smulovitch, Thacker & Woll, 2004; Atkinson, Marshall, Woll & Thacker, 2005). Marshall et al. (2004) investigated naming by a single patient with anomia, finding no difference between accuracy in producing iconic and non-iconic signs. Atkinson et al. (2005) investigated deaf signers with stroke affecting the right or left hemisphere. A sign-picture matching task in which the patients were asked to choose a picture (out of 5) corresponding to an iconic or non-iconic sign showed no difference between iconic and non-iconic signs for patients from either group (left or right hemisphere lesion). Individual patient data was not presented, so we do not know if individual patients would have shown an effect of iconicity. However, at least in BSL, there is no evidence as yet that iconicity confers an advantage following damage to the brain.

1.1 The present study

Here, we report the performance of thirteen individuals with aphasia tested on a number of different tasks comparing iconic and non-iconic words. Assessing whether and in what tasks such effects are observed and whether they correlate with other measures of semantic processing (e.g., imageability, background tests of semantic processing), we can begin to explore the mechanisms behind iconic effects.

We used four tasks that are routine for the assessment of aphasia and rely on phonology-semantics mappings to a greater or lesser extent. For production, we used reading aloud and word repetition. Reading aloud involves semantic activation to a greater extent than word repetition, whilst word repetition places demands on mapping input phonology to articulatory output (e.g. Bates, Burani, D'Amico & Barca, 2001; Bose, Colangelo & Buchanan, 2011; Hickok & Poeppel, 2004; Woollams, Lambon Ralph, Plaut & Patterson, 2007). Therefore, effects of iconicity, if any, should be greater in reading than in repetition. For word recognition, we used auditory and visual lexical decision. We expect auditory lexical decision to place greater demands on mapping phonology to semantics than visual lexical decision as input phonology is directly targeted (e.g. Chen & Cutler, 1997). Therefore, iconicity effects may be greater in the former than in the latter. Across these tasks, we compared accuracy scores for a set of iconic words and a set of matched control words. In addition, we correlated observed effects of iconicity with two canonical measures of semantic processing (word-word / picture-picture association and performance with high vs. low imageability words) administered as background assessments for the patients in this study. This analysis allowed us to assess whether iconicity effects can be reduced to semantic variables alone.

2.0 Methods and materials

2.1 Design

An experimental case-series design was used. Performance (accuracy) for a group of individuals with aphasia was compared within subjects on tasks manipulating iconicity (iconic vs control words). Error types were compared within subjects for speech production tasks (iconic vs control). To evaluate performance

against neurologically healthy adults, individual data was compared to control data using procedures outlined in Crawford & Garthwaite (2005a, 2005b) and Garthwaite & Crawford (2004).

2.2 Participants

2.2.1. Participants with Aphasia

Thirteen participants with aphasia (PWA) took part. All had English as their native language. The group had 6 females and 7 males, with a mean age of 62.5 (SD = 15.3) years, mean education of 13.2 (2.6) years and mean time since stroke of 75.7 (55.3) months. Table 1a presents demographic and background information for all PWA.

All had aphasia as a result of cerebrovascular accident, with a current diagnosis of aphasia based on assessment with the Western Aphasia Battery Revised (WAB-R; Kertesz, 2006) or the Boston Diagnostic Aphasia Examination Third Edition (BDAE-3; Goodglass, Kaplan and Berresi, 2000). Five individuals presented with anomic aphasia, six with Broca's aphasia and two with conduction aphasia. Seven individuals presented with acquired apraxia of speech (AoS), based on clinical assessment of AoS symptoms present in repetition and naming (groping, difficulties with words of increasing length, inconsistent speech errors, difficulty initiating speech). See Table 2 for detail of background assessment.

------------------ Table 1a and 1b here ---------------------------------------

To provide additional information about semantic processing, two further groups of assessments were completed. (1) Assessment of the ability to select appropriate

semantic associates, using the Pyramids and Palm Trees (Howard and Patterson 1992) or Camels and Cactus test (Bozeat et al., 2000) in both written and picture formats. According to the assessment cut-offs, participant P1 was impaired for picture association, and P5 was borderline. P6, P7, P10, P12 and P13 were impaired across both word and picture versions, see Table 2. (2) Performance on a set of high and low imageability words. The motivation for this was two-fold. First, to verify that the experimental tasks (repetition, reading aloud, auditory lexical decision and visual lexical decision) were sensitive to semantic variables for this group of patients. Second, to gain additional information about semantic processing on a patient-by-patient basis. Subtests from the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA; Kay, Coltheart and Lesser, 1992) were used. Subtest 9 is repetition, subtest 31 reading aloud, subtest 5 auditory lexical decision and subtest 25 visual lexical decision. All use the same set of 80 items manipulated across imageability (high versus low) and frequency (high versus low). For the purposes of this study, we are interested in the 40 high and 40 low imageability items (collapsing across frequency). In line with the iconicity tasks (see below) half the participants completed auditory lexical decision (n = 6; P1, P3, P4, P8, P10, P11) and half completed visual lexical decision (n = 7; P2, P5, P6, P7, P9, P12, P13). Results from these tasks can be found in Appendix 2 (reading aloud Table A3, repetition, Table A4, Auditory lexical decision Table A5, Visual lexical decision Table A6). There were significant group effects of imageability for repetition and reading aloud, and trends for effects in the two lexical decision tasks. In all cases, high imageability words were more accurate than low imageability words. Imageability effects (difference between performance on high and low imageability items) were used in correlation analysis with results from the semantic associates and iconicity tasks (see Results).

------------------------ Table 2 here ------------------------

2.2.2 Controls

Fourteen control participants were recruited to be age, sex and education matched to the PWA as a group. All participants had English as their native language, with no history of brain injury or other neurological impairment. The group had 7 females and 7 males, with a mean age of 60.86 (SD = 13.03; $p > 0.4$) and mean years in education of 12.43 (SD = 2.17, $p > 0.4$). Table 1b presents demographic and background information for all Controls. Half the controls completed visual lexical decision (C1-C7) and half completed visual lexical decision (C8-C14), see Table 1b. These sub-samples were matched to their respective patient samples for age ($p > 0.9$ and $p > 0.1$ respectively) and years in education ($p > 0.7$ and $p > 0.3$ respectively).

2.3 Materials

2.3.1 Iconic and Control words

A preliminary internet search was conducted to retrieve iconic words in English. The initial list was then reduced by excluding items for which imageability ratings were not available from two large databases of imageability ratings (Bird, Franklin & Howard, 2001; Cortese & Fuggett, 2004). This produced a list of 40 iconic words (see Appendix 1, Table A1). Matched control words were selected for each iconic word. Matching was checked by-group on the following variables: CELEX

combined written and spoken lemma frequency per million (Baayen, Piepenbrock & Gulikers, 1995), Iconic words mean = 3.22 (SD = 3.83), Control words = 2.70 (3.0), independent t-test p>0.4. Imageability, Iconic words = 4.06 (0.88), Control words = 4.07 (0.88), p>0.9. Age of Acquisition ratings (Kuperman et al., 2012) showed that iconic words were on average acquired 2 years earlier than control words; Iconic words = 6.89 (2.20), Control words = 8.15 (2.90), p<0.05. This is in line with previous literature (Monaghan et al., 2014; Thompson et al., 2012). Length in syllables, Iconic words = 1.13 (0.34), Control words = 1.08 (0.27), p>0.4. Length in phonemes, Iconic words = 3.75 (0.92), Control words = 3.98 (0.86), p>0.2. Length in letters, Iconic words = 4.97 (1.12), Control words = 4.95 (1.08), p>0.9. Phonological neighbourhood density, Iconic words = 18.34 (11.40), Control words = 15.84 (9.70), p>0.3. Phonotactic probability for biphonemes (Vaden, Halpin & Hickok, 2009): stressed word average biphone probabilities for Iconic words = 0.0018 (0.0013), Control words = 0.002 (0.0019), p>0.5; unstressed word average biphone probabilities for Iconic words = 0.0035 (0.0023), Control words = 0.0033 (0.0030), p>0.7. Coltheart's N (Medler & Binder, 2005): number of orthographic neighbours, Iconic words = 5.25 (4.71), Control words = 5.68 (4.85), p>0.5; orthographic neighbourhood frequency, Iconic words = 14.68 (19.29), Control words = 24.97 (41.64), p>0.1. To check further for performance differences across the iconic and control words, we extracted mean reaction times and error rates from the English Lexicon Project (ELP; Balota et al., 2007). Data were extracted for lexical decision (visual) and word naming (reading aloud). One iconic word was not available on the database ('clink') giving 40 data points for control words and 39 for iconic words. There were no significant differences between iconic and control words for reaction times or error rates in either task (see Appendix 1, Table A2).

2.3.2 Rating for iconicity

To check the manipulation of iconicity, 10 native speakers of English (5 female, 5 male, mean age = 35.4 years, SD = 13.8) were provided with the written list of iconic and control words in a random order, and asked to rate each word for iconicity on a 10 point scale. Instructions were as follows: "Some words in English sound like what they mean. These words are usually things like animal sounds (*meow, moo, oink*). However, other words also have this property (usually called onomatopoeia). For example, the sounds made by objects in motion or when they hit other objects (e.g. *whoosh or bang*). We also say that these words are iconic, as there is a relationship between the form of the word (its sound) and what it means. There are other kinds of iconic relationships; for example, words that begin with 'gl', like 'gleam', 'glint' or 'glitter', usually refer to a bright visual property that an object has. In this task, you'll read a list of words and asked to rate how iconic you think the word is. A very iconic word would be something like 'meow' or 'moo' where it is completely clear how the sound relates to the meaning. A not at all iconic word has no relationship between the sound and the meaning, e.g. 'smile' or 'type'. A word in the middle might have some element (e.g. like 'glint' or 'glitter') that is iconic, but it's not as obvious as 'meow' and 'moo'."

No examples given in the instructions were present in the iconic word set. Iconic and control words were rated as significantly different, with iconic words having higher overall ratings (mean = 9.09, SD = 0.54) compared to Control words (mean = 2.65, SD = 1.29), $t(78) = -20.87$, $p<0.05$.

2.4 Procedure

Data were collected by two students (ES and DS) completing dissertation projects at the University of Reading 6 months apart. PWA were tested across a

maximum of three sessions lasting up to two hours. The research was conducted and audio recorded in the PWA's own home or at the University of Reading Speech and Language Therapy clinic rooms. All participants received the same instructions. Regular breaks were scheduled during testing and participants were encouraged to request a break if needed.

Tasks were ordered with a latin-square so that participants did not complete a task with the same items consecutively. Thus, in any session a participant was not presented with the same word list in the same order. For iconicity tasks, ten randomly ordered item lists were created and rotated across participants and tasks. Half the participants completed auditory lexical decision (n = 6; P1, P3, P4, P8, P10, P11) and half completed visual lexical decision (n = 7; P2, P5, P6, P7, P9, P12, P13).

Control participants completed only the iconic/control tasks during one half hour session. All controls completed the repetition and reading aloud tasks. Half the controls completed visual lexical decision (C1-C7) and half completed auditory lexical decision (C8-C14).

2.4.1 Reading Aloud

Participants were informed them they would see some words and would be required to read them aloud. For half the participants (P1, P3, P4, P8, P10, P11), words were presented printed on paper, in lower case, using Arial font, size 14. For the other half (P2,P5,P6,P7,P9,P12,P13) the words were presented on a computer screen, in lower case, using Times New Roman font, size 82. For completion of the PALPA subtest, words were presented according to subtest instructions printed on the standard assessment materials (printed on paper, approximate font size 14). In each case, one word was presented at a time until the participant responded or indicated

that they wanted to pass. Responses were audio recorded for later phonetic transcription.

2.4.2 Repetition

Participants were informed that they would hear words and would be required to repeat them. For both iconic/control and PALPA tasks, words were presented verbally by the experimenter with a flat intonation followed by a pause. A word was repeated as many times as requested by the participant.

2.4.3 Lexical decision

Participants were instructed that they would be shown a word and should make a decision about whether it was a real word or not. They were required to answer 'yes' if it was a real word and 'no' if it was not. Participants were able to respond non-verbally if required (e.g. nodding/shaking their head). If participants changed their answer before the presentation of the next word the second answer was accepted. For manipulations of iconicity, there were 160 items: 80 real words (the 40 iconic and 40 control words) and 80 nonwords. Nonwords were generated from the ARC Nonword Database (Rastle, Harrington & Coltheart, 2002). All nonwords had orthographically existing onsets and bodies, and legal bigrams. They were generated to have the same range (set minimum and maximum) for length in letters, length in phonemes, and phonological neighbourhood density as the control and iconic words.

2.4.3.1 Auditory lexical decision

Words were read with a flat intonation with a pause after each response. A word was repeated as many times as requested by the participant. Participants P1, P3, P4, P8, P10 and P11 completed this task.

2.4.3.2 Visual lexical decision

Words were presented on a laptop computer using a PowerPoint presentation. The words were presented 1 per slide in size 82 lowercase Times New Roman font. Words were presented until the participant responded, or indicated they wanted to pass the trial. Participants P2, P5, P6, P7, P9, P12 and P13 completed this task.

3.0 Analysis

Statistical comparisons were completed using R (R Core Team, 2013). As error data was analysed (averages across 0s and 1s), non-parametric statistics were used; group within subjects comparisons with paired Wilcoxon signed rank tests; and correlations with Spearmen tests. Partial correlations were completed with the pcor.test function (Kim & Yi, 2006; 2007). To compare the difference between iconic and control words for each patient against the difference in those conditions exhibited by controls, we used the Revised Standardized Difference Test (RSDT) (Crawford & Garthwaite, 2005a, 2005b; Garthwaite & Crawford, 2004). This was developed to compare the difference between patient performance on two or more tasks (i.e. testing for dissociation). Here, the two item sets are treated as the two different 'tasks', and the difference in performance between them is evaluated against control data. The RSDT controls for Type I error rates when there are correlations between the tasks under study; we entered simple correlations between raw scores from the control groups.

3.1 Scoring for speech production tasks

If the participant successfully self-corrected their response before the next word was presented the response was scored correct. Errors produced initially during such trials were still included in further analyses. Errors were classified as formal (real word, phonologically related to the target, e.g. 'creak' – 'crack'), phonological (non-word, e.g. 'broil' – 'braw'), semantic (real word, semantically related to the

target, e.g. 'fizz' – 'bottle'), mixed semantic and formal (real word, e.g. 'bathe' – 'bath') or other (e.g. pass or no response) (classification according to Schwartz et al., 2004).

4.0 Results

4.1. Reading Aloud

There was a significant effect of iconicity on overall accuracy in reading aloud for PWA. A higher proportion of iconic words were accurately read aloud (mean % accuracy = 63.65, SD = 10.46) as compared to control words (mean % accuracy = 57.11, SD = 10.18); paired Wilcoxon signed rank test with continuity correction, W = 6.5, p<0.05, 95% CI = 1.5 to 6.65. For individual patients, 6 showed a significantly larger difference between iconic and control words as compared to control data (P1, P2, P4, P6, P8 & P11). All 6 showed an advantage for iconic words over control words. See Figure 1 and Table 3.

-------- Insert Figure 1 here --------

Figure 1: Overall accuracy for individual PWA and the group for reading aloud iconic and control words

----------- Insert Table 3 here -------------

Figure Caption: asterisks for individual patients show those who have a greater difference between iconic and control words when compared against control data. Asterisks for group data indicate a significant difference when iconic and control words are compared within subjects across the whole patient group. Error bars for the group data are one standard error. All p<0.05, two-tailed.

4.1.1. Errors

A higher number of formal errors were made on control words (mean errors = 4.70, SD = 3.25) than on iconic words (mean errors = 3.23, SD = 2.83), paired Wilcoxon signed rank test with continuity correction, W = 1.5, p<0.05, 95% CI = -3.0 to -0.99. There was a trend for a higher number of mixed errors on control words (mean errors = 1.0, SD = 1.08) than on iconic words (mean errors = 0.38, SD = 0.77); W = 6, p=0.10, 95% CI = -2.0 to 3.16. There was no difference in the number of phonological (non-word) or semantic errors. See Table 4.

----------- Insert Table 4 here -------------

4.2 Repetition

There was no significant effect of iconicity on overall accuracy in repetition for the group comparison. Across the group, a similar proportion of iconic words were accurately read aloud (mean % accuracy = 59.23, SD = 10.17) when compared to control words (mean % accuracy = 56.92, SD = 10.50); Wilcoxon test p>0.05. For individual patients, 9 showed a significantly larger difference between iconic and control words as compared to control data. Of these, 7 showed an advantage for iconic words over control words (P2, P4, P7, P8, P9, P10 & P11). One patient showed no numerical difference (P6) but the case comparison statistics were significant as the severity of performance differed for this patient between conditions when compared to control data. One patient showed an advantage for control words over iconic words (P13). See Figure 2 and Table 5.

-------- Insert Figure 2 here --------

Figure 2: Overall accuracy for individual PWA and the group for repeating iconic and control words

----------- Insert Table 5 here -------------

4.2.1 Errors

There was no difference in the number of errors of any kind produced on iconic and control words during repetition. See Table 4.

4.3 Correlation of effects with other semantic tests

For the Spearman correlations, we first computed the difference between iconic and control words (i.e. overall accuracy score for control words subtracted from overall accuracy score for iconic words) for each individual patient, for each speech production task (repetition and reading aloud). These were correlated separately with scores on the picture and word versions of the semantic association tests. There were no significant correlations. These analyses were repeated as partial correlations (Kim & Yi, 2006; 2007) that partialled out the effect of age, years in education and severity (WAB AQ or BDAE percentile). There were no significant correlations.

We computed the difference for high and low imageability words (i.e. overall accuracy for low imageability words subtracted from the score for high imageability words) for each patient, separately for the repetition and reading aloud. Scores were then correlated by task. That is, a correlation of iconicity and imageability effects for repetition, and a correlation of iconicity and imageability effects for reading aloud. There were no significant correlations. These analyses were repeated as partial correlations (Kim & Yi, 2006; 2007) that partialled out the effect of age, years in

education and severity (WAB AQ or BDAE percentile). There were no significant

correlations.

4.4 Auditory lexical decision

There was a significant effect of iconicity in auditory lexical decision for the

group comparison. Iconic words were recognized as real words (mean % accuracy =

91.25, SD = 6.08) more often than control words (mean % accuracy = 84.58, SD =

7.37); W = 21, p<0.05, 95% CI = 3.75 to 9.99. For individual patients, 2 showed a

significantly larger advantage for iconic words than control words (P4 & P8), when

compared to control data. See Figure 3 and Table 6

-------- Insert Figure 3 here --------

Figure 3: Overall accuracy for individual PWA and the group for auditory lexical

decision on iconic and control words

----------- Insert Table 6 here -------------

4.5 Visual lexical decision

In the group comparison, there was a trend for iconic words to be recognized as real

words (mean % accuracy = 90, SD = 4.27) more often than control words (mean %

accuracy = 86.67, SD = 4.69), this did not reach significance; W = 19, p=0.08, 95%

CI = 3.02 to 7.50. For individual patients, 4 showed a significantly larger difference

between iconic and control words as compared to control data. Of these, 3 showed an

advantage for iconic words (P5, P6, & P9) and 1 showed an advantage for control

words (P7). See Figure 4 and Table 7.

-------- Insert Figure 4 here --------

Figure 4: Overall accuracy for individual PWA and the group for visual lexical decision on iconic and control words.

----------- Insert Table 7 here -------------

4.6 Check for Age of Acquisition

As the item sets were not matched for Age of Acquisition (AoA) we completed a check that AoA was not driving the iconicity effects seen for PWA. We took each PWA who showed a significant effect of iconicity in the case comparison (i.e. individual) analyses. We then compared AoA ratings for iconic and control items on which they had responded correctly (independent samples t-test, two tailed). Results can be found in Appendix 3. For reading aloud, P4's effect of iconicity was confounded with AoA. For repetition, P11's effect of iconicity was confounded with differences in AoA. For auditory lexical decision there were no apparent confounds. For written lexical decision, three of the four PWA (P5, P6 and P9) had effects of iconicity that were confounded with AoA differences. The only PWA whose effect of iconicity was not confounded with AoA for written lexical decision (P7) showed a reverse iconicity effect (see Table 7). Annotations have been made on Tables 3, 5 and 7 to show which PWA have effects of iconicity confounded with AoA.

5.0 Discussion

We compared accuracy of production and recognition for iconic and matched control words, in a group of individuals with aphasia. Results were compared across the group, but also using a case-comparison analysis (Crawford & Garthwaite; 2005a, 2005b; Garthwaite & Crawford, 2004) to find out whether these individuals showed a

larger difference between iconic and control words than expected from control participant data.

For speech production tasks, data showed a significant group advantage for iconic words in reading aloud but not for repetition. However, case-comparisons showed that iconic words were produced more accurately for approximately half the participants in both reading aloud and repetition. One of the six (P4) had an effect confounded with AoA differences. Speech error data supported the consistent effect of iconicity in reading aloud, with fewer real-word phonological errors on iconic words. For repetition, individual performance was more variable. Six individuals showed significantly better performance on iconic words in case-comparison analyses, with one participant showing a significant advantage for control words, and one showing an iconicity effect confounded with AoA differences. There was no difference in the frequency of any error type.

If the advantage for iconic words in speech production comes from output processes alone (e.g. iconic words have more robust phonological and articulatory representations), then we might have expected a more similar group pattern for the two tasks. Reading aloud requires the transformation of orthographic to phonological forms, and for reading real words, research suggests that lexico-semantic information is typically accessed (e.g. Woollams et al., 2007; Binder et al., 2005). Repetition requires the mapping of auditory input into articulatory output (e.g. Buchsbaum et al., 2011) with less reliance on semantic processing (e.g. Hickok & Poeppel, 2004). Therefore, in reading aloud, better performance for iconic words may come from stronger links from lexico-semantics to phonology, or from hypothesized links between the word-form and sensory modalities (e.g. including greater involvement of the right hemisphere; Hashimoto et al., 2006; Kanero et al., 2014). However, in

repetition, better performance for iconic words may arise from better phonological processing during (auditory) word recognition. This latter interpretation is supported by the data from the lexical decision tasks.

In word recognition tasks, there was an advantage for iconic over control words in auditory lexical decision. This was driven by a consistent, small advantage across all participants who completed the task. Two individuals showed an advantage in case comparison analyses, and these were not confounded with AoA differences. It may be that reaction time data for healthy controls would show a similar pattern, with faster reaction times in auditory lexical decision for iconic forms (reaction time data from the English Lexicon Project is only available for visual lexical decision). For visual lexical decision, as for repetition, performance was more variable. There was no group effect for iconicity. Case comparison analyses showed that three individuals had a significant advantage for iconic words and for all three this was confounded with AoA. One showed a significant advantage for control words. Therefore, the effects in written lexical decision appear to be driven by AoA rather than iconicity (Nickels & Howard, 1995). It is less clear how iconic forms would confer a processing advantage in visual lexical decision in the first place, unless phonological information is being accessed to support decision making (Westbury, 2005). It is worth noting that no differences were seen in visual lexical decision times for healthy controls for iconic and control words (from the English Lexicon Project, see Methods).

Overall, an advantage for iconic words is seen across both production and recognition tasks for a number of individuals with aphasia in our group. We found more group and individual effects of iconicity in reading aloud and auditory lexical decision. This fits an interpretation in which iconicity in spoken languages is most

influential when a phonology-semantics mapping is in use. That is, when the task has demands that can tap into the referential benefit of iconicity in the word form.

The prediction that iconicity was not reducible to semantic variables was broadly supported. There was no association between effects of imageability and iconicity, nor between effects of iconicity and background semantic tests. Caution is needed when interpreting a null effect and the range of scores on background assessments of semantic association was limited. It is possible that some variation across tasks is due to differences in the aphasia profiles of our participants. In addition, at least for some PWA, the effects of iconicity were confounded with AoA.

As our patients' sample is small, only descriptive analyses of iconicity effects by aphasia type can be carried out (on raw scores). Patients 1 to 5 had anomic aphasia, P7 to P11 Broca's aphasia and P12 and 13 conduction aphasia. Thus, on Figures 1 and 2, patient type can be scanned along the x-axis. For iconicity and reading aloud (Figure 1a), a higher score on iconic items are present for 3/5 of those with anomic aphasia, 3/6 of those with Broca's aphasia, and 2/2 with conduction aphasia. So for reading aloud there is no clear modulation of the iconicity effect by aphasia type[1]. For iconicity and repetition (Figure 2a), a higher score on iconic items

---

[1] We did not conduct a specific screening for acquired dyslexia, which could have affected patient performance on written lexical decision and reading aloud. One patient (P6) showed a very low score on the written version of the semantic association tests (20% correct) relative to the picture version (65.6% correct), and a low score on reading aloud (30% for iconic and 17.5% for control words). This could indicate the presence of an acquired dyslexia. However, performance on written lexical decision was relatively high (80% for iconic and 77.5% for control words). It

are present for 2/5 of those with anomic aphasia, 5/6 of those with Broca's aphasia, and 1/2 with conduction aphasia. Thus in repetition it is more likely for individuals with Broca's aphasia to show an effect of iconicity (with the caveat that the group analysis for this was not significant, as absolute differences are small as well as variable). Patients 4, 5, 8, 9, 10, 11 and 13 showed symptoms of acquired apraxia of speech. Of these, 4/7 showed a higher score on iconic words during reading aloud, and 5/7 during repetition. Thus, there is a slightly more consistent pattern of iconicity effects in repetition, mirroring the data above for individuals with Broca's aphasia.

To fully explore the relationship between individual patient profiles and effects of iconicity, a larger sample of individuals would be needed. In addition, more challenging assessments of semantic processing (e.g. synonym judgement) and more detailed assessment of phonological processing in both input and output tasks are needed. If the effects of iconicity do lie at the point where form and meaning are mapped onto each other, tasks which tap into the ventral stream of processing (Hickok & Poeppel, 2004) may show more consistent effects and more consistent relationships with iconicity.

We matched the item sets on a number of different variables, but the iconic items had lower age of acquisition ratings (AoA) than the control set. Previous studies have found that iconic words are acquired earlier (e.g. Laing, 2014; Monaghan, Shillcock, Christiansen & Kirby, 2014; Thompson, Vinson, Woll & Vigliocco, 2012) making it a challenge to match item sets on this variable. We completed a post-hoc check that age of acquisition effects were not driving iconicity effects, and found that

is also worth noting that this patient was overall more severe (Aphasia Quotient of 48).

AoA was a probable confound only for a minority of cases (see Appendix 3). Unfortunately we did not collect data to check whether patients showed an AoA effect on a separate set of items, as we did for imageability, in order to compare these effects. In future studies, it may be more useful to select a large set of items (e.g. more than 100) that vary in iconicity so that a regression analysis can be performed by-patient, in order to explore the effect of iconicity and account for AoA.

4.1 Theoretical implications of our findings

Our data argue against views in which iconicity would act to strengthen the activation of phonology or semantics alone. If iconicity led to greater activation of phonological information we should have seen similar results across auditory lexical decision and repetition. Instead, we found the most consistent effects in those tasks that placed greater demands on mapping between semantics and phonology (reading aloud and auditory lexical decision). If iconicity led to greater activation of semantic information, then we would expect to see significant correlations with measures of semantics (i.e. by-patient imageability effects and picture-association scores).

Thus, our results point to iconicity as providing an advantage in the *mapping* of phonological information and properties of semantics. There are two alternative (non-mutually exclusive) ways in which this could be achieved. In a theoretical framework in which there is an identifiable locus for representing semantic information (e.g. Patterson, Nestor & Rogers, 2007), words with iconic properties may enjoy additional connections from the semantic system to modality-specific features. In this case, iconic forms would be more robust than non-iconic forms to neurological damage as they are represented with greater redundancy within the language system itself. That is, they are protected in a similar fashion to high

frequency, high imageability or low age-of-acquisition words (Bird, Franklin & Howard, 2001).

On an alternative account, more iconic properties are represented by direct connections between the phonological form and modality-specific information. This could be part of a distributed network of language processing (e.g. Skipper & Small, 2005) and in line with embodied theories of semantics (Meteyard et al., 2012), in which iconic forms have an extra route to activate experience. This latter possibility is in line with the proposal by Kanero et al. (2014) according to which iconic forms would be processed as both linguistic forms and as sound-symbols (Kanero et al., 2014). Here, iconic forms would be more robust to neurological damage because they are additionally processed by areas engaged in cross-modal integration, including right hemisphere brain regions (Westbury, 2005; Hashimoto et al., 2006; Kanero et al., 2014). Therefore, when left-hemisphere language networks are lesioned, iconic forms are retained as there are undamaged pathways available to process them.

We briefly reviewed the advantage for iconic words during vocabulary learning, in both children and adults (Imai, Kita, Nagumo & Okada, 2008; Kantartzis, Imai & Kita, 2011; Kovic, Plunkett & Westermann, 2010; Nygaard, Cook and Namy; 2009). Our data provides a complement to these findings. It may be the case that effects of iconicity for lexical items are limited in the healthy, mature language system (e.g. Westbury, 2005), but the referential benefit of iconicity appears when the language system is challenged. That is, when new lexical items are being acquired, or existing lexical items are more difficult to retrieve. Future research should explore the potential of a therapeutic benefit from iconic forms. It is possible that, as for learning manual gestures (Coelho & Duffy, 1986), individuals with aphasia may be better able to learn successful production and comprehension of word-forms that are iconic.

4.2 Iconicity and Arbitrariness

If iconicity is beneficial in language learning and in aphasia, why do we have so little of it in languages such as English? One possibility is that while iconicity is useful and therefore it is deployed in languages in order to map linguistic forms (spoken or signed) onto properties of world-referents, arbitrariness would subserve a different function, namely to make the linguistic signal more easily discriminable. The argument here is that if all words referring to semantically similar objects were also to be phonologically similar, it would be more difficult to learn words as well as to process them, for example in noisy situations (Monaghan, Christiansen & Fitneva, 2011; Perniss et al., 2010). Moreover, if we consider communication more broadly, it is clearly the case that iconicity is present not only in the phonological form of words but also in the prosody as well as in the gestures that accompany and are time-locked with speech (Skipper et al., 2009). These gestures often iconically represent what is being said and are integrated with speech during language comprehension (Kelly, Özyürek & Maris, 2009) and strong functional connectivity has been shown between motor planning areas and anterior semantic areas suggesting tuning of brain responses to take into account both the semantic content of the speech as well as of the gestures (Skipper et al., 2009). Individuals with aphasia have been shown to produce spontaneous iconic gestures during conversation (Caldognetto & Poggi, 1995; Wilkinson, 2013) and when experiencing word-finding difficulties (Cocks, Dipper, Middleton & Morgan, 2011). They are also better able to learn gestures with high rather than low levels of iconicity (Coelho & Duffy, 1986). Thus, because iconicity is beneficial it would be retained across languages, although in spoken languages it

would be primarily expressed by co-speech gestures (see Perniss & Vigliocco, in press).

In conclusion, we have presented data showing a benefit of iconic word forms in a mixed group of individuals with aphasia. More consistent effects were found in reading aloud and auditory lexical decision. Iconic forms are argued to give a referential benefit, providing stronger links between phonology and semantics. The data presented here is complementary to existing research showing that iconic mappings are beneficial for word learning. Here, we provide some preliminary evidence that iconic forms are also more robust to neurological damage. As such, there may be potential for using iconic word-forms to support language rehabilitation in aphasia.

Acknowledgements

References

Atkinson, J., Marshall, J., Woll, B., & Thacker, A. (2005). Testing comprehension abilities in users of British Sign Language following CVA. Brain and language, 94(2), 233-248.

Baayen, H.R., Piepenbrock, R. & Gulikers, L. (1995) The CELEX English Lexical Database. Release 2 (CD-ROM).

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English Lexicon Project. Behavior Research Methods, 39, 445-459.

Bates, E., Burani, C., D'Amico, S., & Barca, L. (2001). Word reading and picture naming in Italian. Memory & Cognition, 29(7), 986-999.

Beattie, G., & Shovelton, H. (2002). What properties of talk are associated with the generation of spontaneous iconic hand gestures?. British Journal of Social Psychology, 41(3), 403-417.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebral Cortex, 19(12), 2767-2796.

Binder, J. R., Medler, D. A., Desai, R., Conant, L. L., & Liebenthal, E. (2005). Some neurophysiological constraints on models of word naming. Neuroimage, 27(3), 677-693.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*(1), 73-79.

Bose, A., Colangelo, A. and Buchanan, L. (2011) Effect of phonetic complexity on word reading and repetition in deep dyslexia. Journal of Neurolinguistics, 24 (4). pp. 435-444.

Botha, R. (2007). On homesign systems as a potential window on language evolution. Language and Communication, 27, 41-53.

Bozeat S, Lambon Ralph MA, Graham KS, Patterson K, Wilkin H, Rowland J, et al. A duck with four legs: investigating the structure of conceptual knowledge using picture drawing in semantic dementia. Cogn Neuropsychol 2003; 20: 27–47.

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory–an aggregate analysis of lesion and fMRI data. Brain and language, 119(3), 119-128.

Burke, D.M. & Shafto, M.A. (2008) Language and Ageing. In Craikk, F.I.M. & Salthouse, T.A. (Eds) The Handbook of Aging and Cognition. New York and Hove: Psychology Press.

Caldognetto, E. M., & Poggi, I. (1995). Creative iconic gestures: some evidence from Aphasics. Iconicity in language, 110, 257.

Chen, H. C., & Cutler, A. (1997). Auditory priming in spoken and printed word recognition; in Chen, H-C (Ed) The cognitive processing of Chinese and related Asian languages, Ch. 5, 77-81. Chinese University Press, Hong Kong.

Cocks, N., Dipper, L., Middleton, R., & Morgan, G. (2011). What can iconic gestures tell us about the language system? A case of conduction aphasia. International Journal of Language & Communication Disorders, 46(4), 423-436.

Coelho, C.A. & Duffy, R.J. (1986) Effects of iconicity, motoric complexity and linguistic function on sign acquisition in severe aphasia. Perceptual and Motor Skills, 63, 519-530.

Connell, L., & Lynott, D. (2014). I see/hear what you mean: Semantic activation in visual word recognition depends on perceptual attention. Journal of Experimental Psychology: General, 143(2), 527.

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 384-387.

Crawford, J. R., & Garthwaite, P. H. (2005a). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: Evaluation of alternatives using Monte Carlo simulations and revised tests for dissociations. Neuropsychology, 19, 318-331.

Crawford, J. R. & Garthwaite, P. H. (2005b) Evaluation of criteria for classical dissociations in single-case studies by Monte Carlo simulation. Neuropsychology, 19, 664-678.

Firth, J.R. (1930) Speech. London, Benn's Sixpenny Library. Reprinted in Stevens, P. (Ed.) (1964) The Tongues of Men and Speech. London: Oxford University Press.

Fischer, O. & Nänny, M. (1999). Introduction: Iconicity as a creative force in language use. In Nänny, M. & Fischer, O. (Eds.), Form Miming Meaning Iconicity in Language and Literature (pp 15-36). Amsterdam: John Benjamins Publishing Co.

Garthwaite, P. H., & Crawford, J. R. (2004). The distribution of the difference between two t-variates. Biometrika, 91, 987-994.

Goldin-Meadow, S. (2005). The resilience of language: What gesture creation in deaf children can tell us about how all children learn language. Psychology Press.

Goodglass, H., Kaplan, E., & Barresi, B. (2000). *Boston Diagnosis Aphasia Examination – Third Edition.* Austin: PRO-ED.

Hamano, S. (1998). The Sound-Symbolic System of Japanese. Center for the Study of Language and Information, Stanford University, Stanford, CA,

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. Psychological review, 111(3), 662.

Hashimoto, T., Usui, N., Taira, M., Nose, I., Haji, T., & Kojima, S. (2006). The neural mechanism associated with the processing of onomatopoeic sounds. Neuroimage, 31(4), 1762-1770.

Hickok, G. & Poeppel, D. (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition, 92, 67-99.

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

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. Journal of cognitive neuroscience, 19(7), 1175-1192.

Howard, D., & Patterson, K. E. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. Thames Valley Test Company.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. Cognition, 109(1), 54-65.

Kanero, J., Imai, M., Okuda, J., Okada, H. & Matsuda, T. (2014) How Sound Symbolism Is Processed in the Brain: A Study on Japanese Mimetic Words. PLOS ONE, 9(5), e97905.

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound‑symbolism facilitates word learning in English‑speaking children. Cognitive Science, 35(3), 575-586.

Kay, J., Lesser, R. & Coltheart, M. (1992). *Psycholinguistic assessments of language processing in aphasia (PALPA).* Lawrence Erlbaum, London.

Kelly, S. D., Özyürek, A., & Maris, E. (2009). Two Sides of the Same Coin Speech and Gesture Mutually Interact to Enhance Comprehension. Psychological Science.

Kertesz, A. (2007). *Western Aphasia Battery Revised.* San Antonio: PsychCorp.

Kim, S.H. and Yi, S. (2006) Correlated asymmetry between sequence and functional divergence of duplicate proteins in Saccharomyces cerevisiae. Molecular Biology and Evolution, 23: 1068.

Kim, S.H. and Yi, S. (2007) Understanding relationship between sequence and functional evolution in yeast proteins . Genetica, 131: 151.

Kovic, V., Plunkett, K. & Westermann, G. (2010). The shape of words in the brain. Cognition, 114 (1), 19-28.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. Behavior Research Methods, 44(4), 978-990.

Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. First Language, 34(5), 387-405.

Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. Cognition, 42(1), 1-22.

Marshall, J., Atkinson, J., Smulovitch, E., Thacker, A., & Woll, B. (2004). Aphasia in a user of British Sign Language: Dissociation between sign and gesture. Cognitive neuropsychology, 21(5), 537-554.

Medler, D.A., & Binder, J.R. (2005). MCWord: An On-Line Orthographic Database of the English Language. http://www.neuro.mcw.edu/mcword/

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. Cortex, 48(7), 788-804.

Miller, G. A., & Johnson-Laird, P. N. (1976). Language and perception. Belknap Press. Cambridge, UK.

Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. Journal of Experimental Psychology: General, 140(3), 325.

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language. Philosophical Transactions of the Royal Society B.

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? Neuropsychologia, 33, 1281-1303.

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. Cognition, 112(1), 181-186.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. Nature Reviews Neuroscience, 8(12), 976-987.

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. Frontiers in psychology, 1, 227.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. NeuroImage, 62, 816-847.

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The Effects of Feature‑Label‑Order and Their Implications for Symbolic Learning. Cognitive Science, 34(6), 909-957.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought and language. Journal of Consciousness Studies, 8, 3–34.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. Quarterly Journal of Experimental Psychology, 55A, 1339-1362.

Reilly, J., Cross, K., Troiani, V., & Grossman, M. (2007). Single‑word semantic judgements in semantic dementia: Do phonology and grammatical class count? Aphasiology, 21(6-8), 558-569. doi:10.1080/02687030701191986

Reilly, J., & Kean, J. (2007). Formal distinctiveness of high- and low-imageability nouns: analyses and theoretical implications. Cognitive Science, 31(1), 157-68. doi:10.1080/03640210709336988

Reilly, J., Westbury, C., Kean, J., & Peelle, J. E. (2012). Arbitrary symbolism in natural language revisited: when word forms carry meaning. PLoS One, 7(8), e42286. doi:10.1371/journal.pone.0042286

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. Brain and language, 128(1), 18-24.

de Saussure, F. (1916) Course in General Linguistics. McGraw-Hill, New York.

Shewan, C. M., & Kertesz, A. (1980). Reliability and validity characteristics of the Western Aphasia Battery (WAB). Journal of Speech and Hearing Disorders, 45(3), 308-324.

Schmidtke, D. S., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. Frontiers in psychology, 5, 80.

Schwartz, M., Wilshire, C., Gagnon, D. & Polansky, M. (2004). Origins of Nonword Phonological Errors in Aphasic Picture Naming. *Cognitive Neuropsychology 21* (2/3/4) 159-186

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2009). Gestures orchestrate brain networks for language understanding. Current Biology, 19(8), 661-667.

Skipper, J., & Small, S. (2005). fMRI studies of language. The Encyclopedia of Language & Linguistics, 496-511.

Stanovich, K.E. & Baur, D.W (1978) Experiments on the spelling-to-sound regularity effect in word recognition. Memory & Cognition, 6(4), 410-415.

Taub, S. (2001). Language from the body. Iconicity and Meatphor in American Sign Language. Cambridge University Press, Cambridge, UK.

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British sign language: effects of iconicity for phonological decisions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36(4), 1017.

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.

Vaden, K.I., Halpin, H.R., & Hickok, G.S. (2009). Irvine Phonotactic Online Dictionary, Version 2.0. [Data file]. Available from http://www.iphod.com.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. Brain and Language, 93(1), 10-19.

Westbury, C., & Moroschan, G. (2009). Imageability x phonology interactions in lexical access. The Mental Lexicon, 4(1), 115-145.

Wilkinson, R. (2013) Gestural Depiction in Acquired Language Disorders: On the Form and Use of Iconic Gestures in Aphasic Talk-In-Interaction. Augmentative and Alternative Communication, 29(1), 68-82.

Wilkinson, R., Beeke, S., & Maxim, J. (2010). Formulating actions and events with limited linguistic resources: Enactment and iconicity in agrammatic aphasic talk. Research on language and social interaction, 43(1), 57-84.

Woollams, A. M., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. (2007). SD-squared: on the association between semantic dementia and surface dyslexia. Psychological review, 114(2), 316.

Yap, D. F., So, W. C., Melvin Yap, J. M., Tan, Y. Q., & Teoh, R. L. S. (2011). Iconic gestures prime words. Cognitive science, 35(1), 171-183.

Table 1a and 1b: Demographic and background information for participants with aphasia and control participants

ID	Gender	Age (years)	Education (years)	Time since onset (months)
P1	F	78	11	15
P2	M	47	11	153
P3	M	73	11	45
P4	F	58	14	22
P5	M	69	18.5	166
P6	F	76	10	132
P7	M	72	13	53
P8	F	67	13	140
P9	M	31	16	16
P10	M	73	12	68
P11	M	76	12	55
P12	F	42	13	54
P13	F	50	17	65
Mean		62.5	13.2	75.7
SD		15.3	2.6	55.3

ID	Gender	Age (years)	Education (years)
C1	F	30	14
C2	F	75	10
C3	M	71	11
C4	M	49	14
C5	F	50	11
C6	M	44	18
C7	M	70	15
C8	F	64	11
C9	M	66	11
C10	F	74	13
C11	M	70	12
C12	F	63	12
C13	F	60	11
C14	M	66	11
Mean		60.9	12.4
SD		13.0	2.2

Table 2: Summary data from background assessment of aphasia.

ID	Aphasia Type[a]	Severity[a]	Apraxia of speech[b]	Spontaneous / simple social speech[c]	Object / picture naming[c]	Responsive speech / naming[c]	Repetition (word)[c]	Auditory word comp. / rec.[c]	Following commands[c]	Semantic Assocoation[d]	
										Words	Pictures
P1	Anomic	88 AQ	N	18/20	51/60	10/10	98/100	60/60	56/80	92.3	88.5
P2	Anomic	91 AQ	N	19/20	60/60	10/10	84/100	60/60	60/80	90.6	96.9
P3	Anomic	91 AQ	N	18/20	58/60	10/10	82/100	59/60	80/80	96.2	98.1
P4	Anomic	92 AQ	N	19/20	56/60	9/10	86/100	60/60	76/80	94.2	98.1
P5	Anomic	88 AQ	Y	17/20	52/60	10/10	86/100	60/60	80/80	93.8	87.5
P6	Broca's	48 AQ	Y	9/20	46/60	6/10	24/100	54/60	34/80	20.3	65.6
P7	Broca's	69 AQ	N	13/20	47/60	8/10	63/100	58/60	55/80	85.9	85.9
P8	Broca's	26 AQ	N	6/20	0/60	0/10	0/100	51/60	40/80	96.2	92.3
P9	Broca's	40th %	Y	7/7	2/15	5/10	3/5	11/16	7/10	92.2	95.3
P10	Broca's	37 AQ	Y	3/20	37/60	8/10	29/100	58/100	39/80	89.4	76.9
P11	Broca's	50 AQ	Y	7/20	44/60	7/10	36/100	59/60	44/80	96.2	98.1
P12	Conduction	100th %	Y	7/7	15/15	8/10	5/5	15/16	9/10	84.4	76.6
P13	Conduction	48 AQ	N	8/20	21/60	8/10	38/100	52/60	65/80	79.7	71.9

a From Boston Diagnostic Aphasia Examination – 3rd Edition (BDAE; Goodglass, Kaplan and Berresi, 2000) or Western Aphasia Battery Revised (WAB-R; Kertesz, 2006). The WAB-R provides an Aphasia Quotient (AQ) and the BDAE provides an aphasia severity score which links to a percentile (%). For both, a higher value indicates less severe aphasia / better language function. b From informal assessment of repetition and naming. c raw scores on subtests taken from the from WAB-R or BDAE. Auditory word comp. / rec. = auditory word comprehension / recognition. d From Camels & Cactus (CCT, Bozeat et al., 2000) or Pyramids & Palm Trees (PPT, Howard and Patterson 1992). Presented as proportion correct. Clinical cut off for the CCT is 87.5%, and for the PPT 90%. Individuals with impairments according to these cut-offs are shaded.

Table 3: Case comparison data for patients on reading aloud

Patient	Control sample Mean (SD)			Case's score		Sig. test*		Estimated % of control population exhibiting a difference more extreme than the case	Effect size (Z-DCC) for the difference between case and controls (95% Bayesian Credible Interval)	
	n	Iconic	Control	Iconic	Control	t	p	Point	Point	(95% CI)
	14	98.75 (2.73)	97.5 (2.59)							
P1				100	87.5	3.76	0.01	0.12	4.21	2.51 6.23
P2				92.5	77.5	4.67	0.01	0.02	5.29	2.46 8.66
P3				75	75	0.02	0.98	49.10	-0.025	-4.07 4.03
P4				87.5	72.5	4.75	0.01[a]	0.02	5.384	1.88 9.94
P5				90	90	0.28	0.78	39.24	-0.31	-1.83 1.18
P6				30	17.5	4.87	0.01	0.02	5.52	-7.21 19.02
P7				97.5	95	0.45	0.66	22.03	0.49	-0.14 1.16
P8				7.5	2.5	2.81	0.01	0.74	3.121	-12.88 19.58
P9				80	82.5	0.96	0.35	17.64	-1.06	-4.12 1.87
P10				5	10	0.54	0.60	29.88	-0.594	-16.33 15.10
P11				72.5	45	8.58	0.01	0.00	10.36	3.04 18.91
P12				87.5	87.5	0.24	0.82	40.84	-0.26	-2.19 1.64
P13				2.5	0	2.06	0.06	3.01	2.27	-14.41 19.30
Mean				63.65 (37.72)	57.12 (36.70)					

*two-tailed   [a] effect confounded with differences in age of acquisition, see section 4.6 in Results and Appendix 3.

Table 4: Mean (Standard deviation) for errors on iconic and control items during repetition and reading aloud.

Task and condition	Error Type			
	Formal	Phonological	Mixed	Semantic
Repetition				
Iconic	6.31 (5.23)	7.46 (9.47)	0.23 (0.60)	0.0 (0)
Control	5.69 (5.68)	8.38 (11.13)	0.38 (0.65)	0.08 (0.28)
Reading Aloud				
Iconic	3.23 (2.83)	6.31 (9.30)	0.38 (0.77)	0.28 (0.28)
Control	4.69 (3.25)	6.77 (9.26)	1.0 (1.08)	0.15 (0.55)

Table 5: Case comparison data for patients on repetition

Patient	Control sample Mean (SD)			Case's score		Sig. test*		Estimated % of control population exhibiting a difference more extreme than the case	Effect size (Z-DCC) for the difference between case and controls (95% Bayesian Credible Interval)	
	n	Iconic	Control	Iconic	Control	t	p	Point	Point	(95% CI)
	14	97.32 (3.46)	97.32 (2.49)							
P1				95	100	1.91	0.08	3.91	-2.13	-3.17 -1.23
P2				95	90	2.57	0.03	1.40	2.76	1.40 4.35
P3				90	95	1.30	0.22	10.78	-1.44	-2.57 -0.44
P4				90	85	3.07	0.01	0.45	3.45	1.37 5.84
P5				90	90	0.91	0.38	19.10	1.00	-0.35 2.45
P6				22.5	22.5	8.30	0.01	0.00	10.25	-2.49 24.02
P7				55	50	6.89	0.01	0.00	8.24	0.37 16.93
P8				5	0	11.23	0.01	0.00	15.09	-1.26 32.95
P9				22.5	12.5	11.33	0.01	0.00	15.14	1.07 30.70
P10				35	27.5	9.57	0.01	0.00	12.21	0.60 25.02
P11				77.5	72.5	4.51	0.01[a]	0.03	5.17	1.06 9.76
P12				90	87.5	2.00	0.07	3.35	2.23	0.54 4.12
P13				2.5	7.5	8.49	0.01	0.00	10.54	-4.99 27.16
Mean				59.23 (10.17)	56.92 (10.50)					

*two-tailed   [a] effect confounded with differences in age of acquisition, see section 4.6 in Results and Appendix 3.

Table 6: Case comparison data for patients on auditory lexical decision

Patient	Control sample Mean (SD)			Case's score		Sig. test*		Estimated % of control population exhibiting a difference more extreme than the case	Effect size (Z-DCC) for the difference between case and controls (95% Bayesian Credible Interval)	
	n	Iconic	Control	Iconic	Control	t	p	Point	Point	(95% CI)
	7	99.29 (1.22)	98.21 (0.44)							
P1				100	97.5	0.65	0.54	26.93	0.80	-0.04 1.71
P3				87.5	80	2.35	0.06	2.87	-2.94	-9.02 2.32
P4				100	90	2.65	0.03	1.91	3.34	1.46 5.76
P8				62.5	50	7.72	0.01	0.01	-12.10	-30.87 3.31
P10				97.5	95	0.24	0.81	40.95	-0.29	-1.47 0.83
P11				100	95	1.34	0.23	11.47	1.64	0.55 2.94
Mean				91.25 (6.08)	84.58 (7.37)					

*two-tailed

Table 7: Case comparison data for patients on visual lexical decision

Patient	Control sample Mean (SD)			Case's score		Sig. test*		Estimated % of control population exhibiting a difference more extreme than the case	Effect size (Z-DCC) for the difference between case and controls (95% Bayesian Credible Interval)	
	n	Iconic	Control	Iconic	Control	t	p	Point	Point	(95% CI)
	7	99.29 (1.22)	96.07 (6.27)							
P2				95	85	1.68	0.14	7.18	-2.13	-4.67 0.00
P5				72.5	67.5	10.59	0.01[a]	0.00	-21.16	-38.17 -8.10
P6				80	77.5	8.73	0.01[a]	0.01	-15.62	-28.01 -6.13
P7				92.5	95	4.63	0.01	0.18	-6.56	-11.35 -2.89
P9				77.5	75	9.44	0.01[a]	0.00	-17.63	-31.61 -6.92
P12				100	95	0.74	0.49	24.37	0.92	0.06 1.88
P13				100	100	0.04	0.97	48.44	-0.05	-0.91 0.81
Mean				90 (4.27)	86.67 (4.69)					

*two-tailed   [a] effect confounded with differences in age of acquisition, see section 4.6 in Results and Appendix 3.

Figure Caption: asterisks for individual patients show those who have a greater difference between iconic and control words when compared against control data. Asterisks for group data indicate a significant difference when iconic and control words are compared within subjects across the whole patient group. Error bars for the group data are one standard error. All p<0.05, two-tailed.

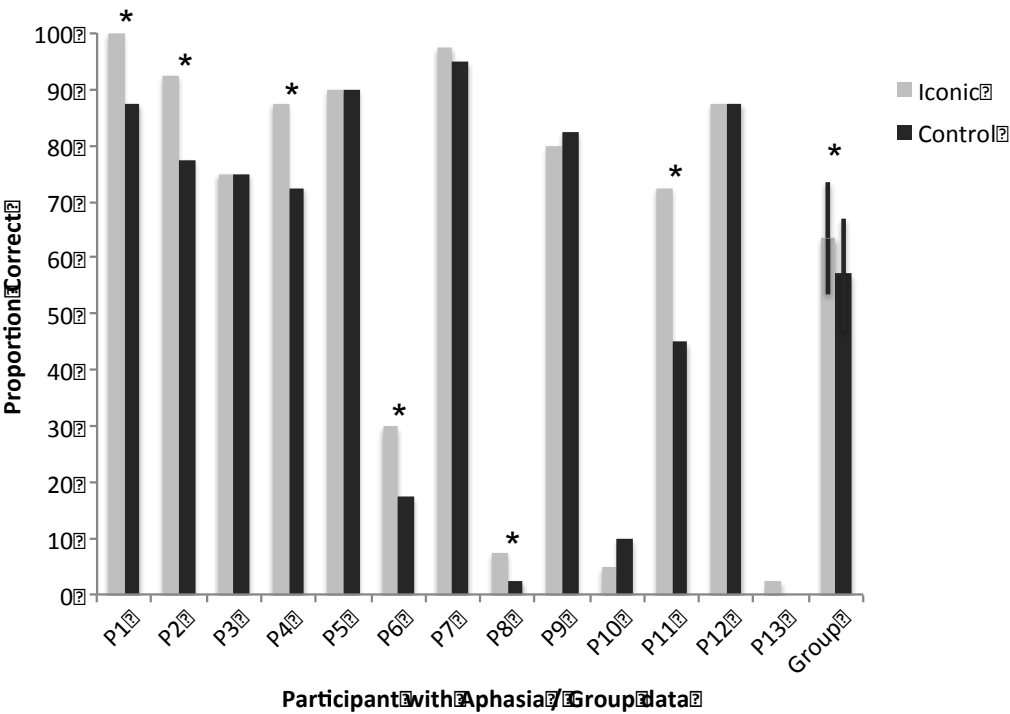Figure 1: Overall accuracy for individual PWA and the group for reading aloud iconic and control words

Figure 2: Overall accuracy for individual PWA and the group for repeating iconic and control words
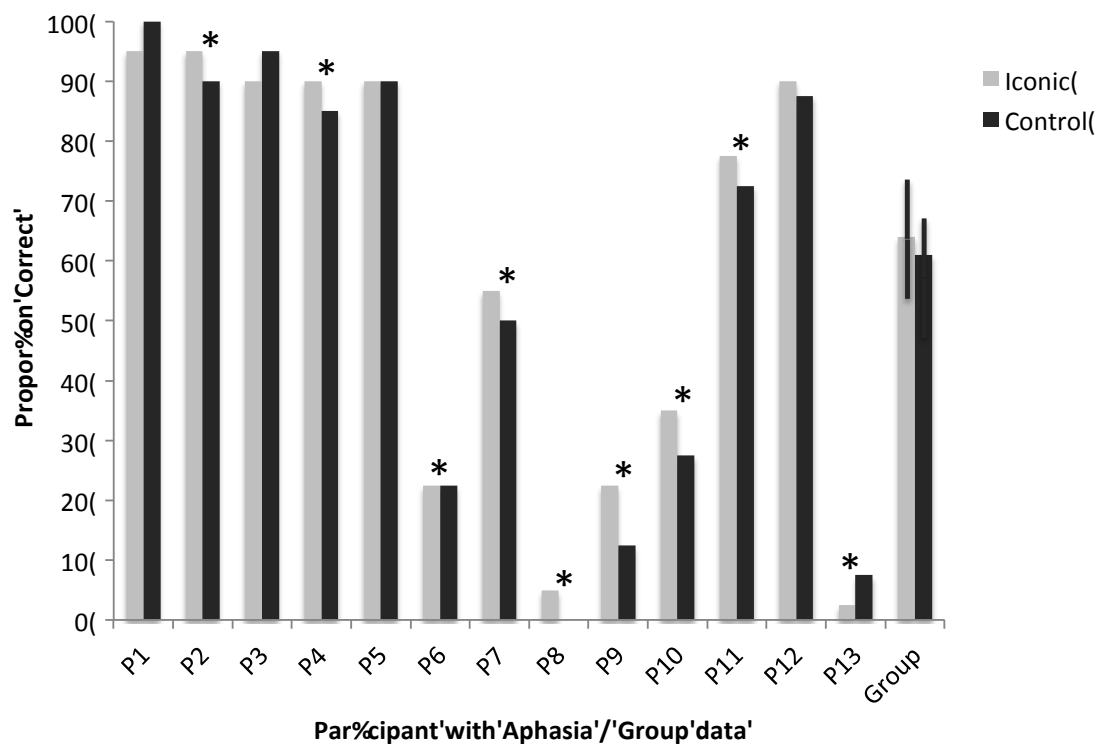
Figure 3: Overall accuracy for individual PWA and the group for auditory lexical decision on iconic and control words
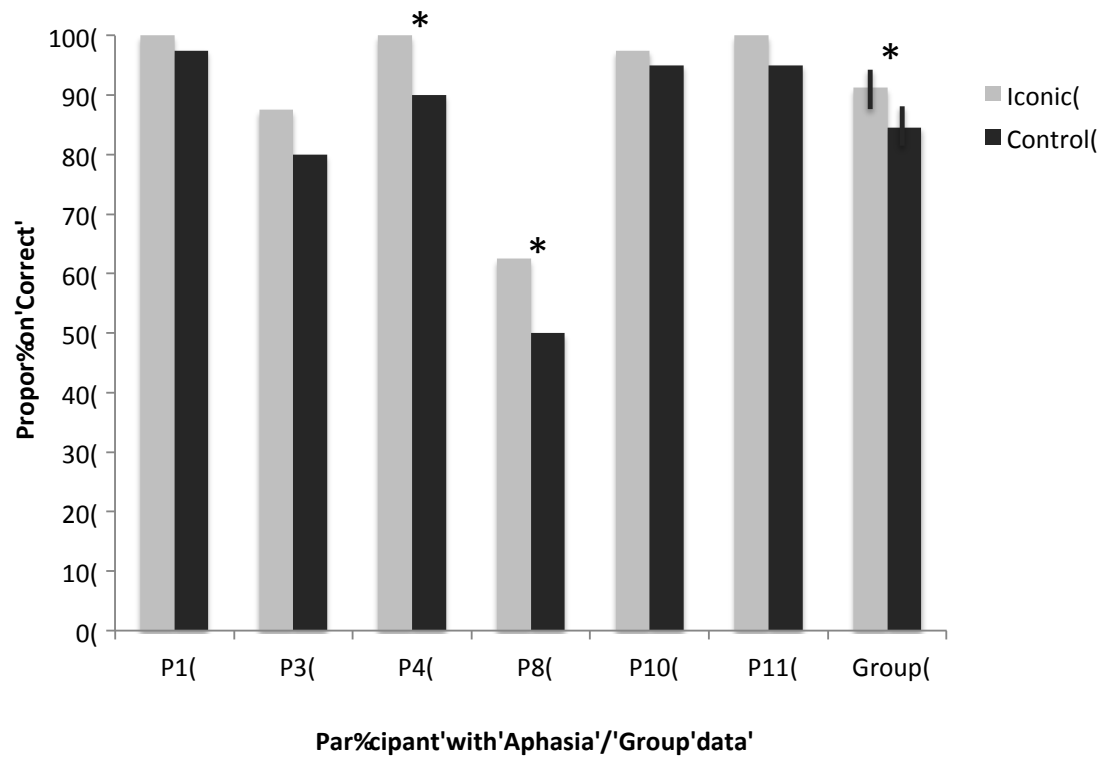
Figure 4: Overall accuracy for individual PWA and the group for visual lexical decision on iconic and control words