Received: 12 April 2019 | Revised: 21 July 2019 | Accepted: 31 July 2019
DOI: 10.1002/hbm.24772

RESEARCH ARTICLE

WILEY

Affective iconic words benefit from additional sound–meaning integration in the left amygdala

Arash Aryani[1] | Chun-Ting Hsu[2] | Arthur M. Jacobs[1,3]

[1]Department of Experimental and Neurocognitive Psychology, Freie Universität Berlin, Germany
[2]Kokoro Research Center, Kyoto University, Kyoto, Japan
[3]Centre for Cognitive Neuroscience Berlin (CCNB), Berlin, Germany

Correspondence
Arash Aryani, Department of Experimental and Neurocognitive Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, D-14195 Berlin, Germany.
Email: arash.aryani@fu-berlin.de

Funding information
German Research Foundation

Abstract
Recent studies have shown that a similarity between sound and meaning of a word (i.e., iconicity) can help more readily access the meaning of that word, but the neural mechanisms underlying this beneficial role of iconicity in semantic processing remain largely unknown. In an fMRI study, we focused on the affective domain and examined whether affective iconic words (e.g., high arousal in both sound and meaning) activate additional brain regions that integrate emotional information from different domains (i.e., sound and meaning). In line with our hypothesis, affective iconic words, compared to their non-iconic counterparts, elicited additional BOLD responses in the left amygdala known for its role in multimodal representation of emotions. Functional connectivity analyses revealed that the observed amygdalar activity was modulated by an interaction of iconic condition and activations in two hubs representative for processing sound (left superior temporal gyrus) and meaning (left inferior frontal gyrus) of words. These results provide a neural explanation for the facilitative role of iconicity in language processing and indicate that language users are sensitive to the interaction between sound and meaning aspect of words, suggesting the existence of iconicity as a general property of human language.

KEYWORDS
affective iconicity, fMRI, language and emotion, left amygdala, neurocognitive poetics, phonaesthetics, sound symbolism

1 | INTRODUCTION

The evolutionary jump from the use of inherent and motivated signs (e.g., a cave painting of a horse representing a horse) toward building unmotivated and arbitrary signs (e.g., using the word "horse" to represent a horse) has been suggested to lay the groundwork for why humans have language (Deacon, 1997). Thus, the arbitrariness of linguistic sign is considered one of the most fundamental properties that grants human language its compositional power, referential flexibility, and productivity (De Saussure, 2011; Gasser, 2004; Hockett, 1958; Monaghan, Christiansen, & Fitneva, 2011), setting humans apart in the animal kingdom by means of a remarkably unique communication system.

However, in contrast to the notion of the absolute arbitrariness, recent empirical data suggest a stand-alone role of sound in meaning-making beyond arbitrary and conventional links (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Perniss, Thompson, & Vigliocco, 2010; Schmidtke, Conrad, & Jacobs, 2014). A prominent type of such nonarbitrary mapping between sound and meaning is iconicity in which sound imitates or resembles some aspects of the meaning. Onomatopoeia (e.g., "click") and ideophones (e.g., "zigzag") represent most instructive examples for this category of words that can evoke sensory, motor, or affective experiences. Contrary to the Saussurean assumption of onomatopoeia and ideophones being a marginal case in language practice, this type of words are nowadays recognized as a widespread phenomenon (Jakobson & Waugh, 1979) forming a major word class equivalent to nouns and verbs in many languages of the world (Dingemanse, 2018).

Iconic words have been suggested to be capable of directly activating the semantic domain that they refer to by bridging the gap between linguistic form and human (sensory, motor and affective) experience (Aryani, Conrad, Schmidtke, & Jacobs, 2018; Perniss & Vigliocco, 2014; Vinson, Thompson, Skinner, & Vigliocco, 2015). Thus, iconicity may provide additional mechanisms for both vocabulary learning and language processing by means of direct sound-meaning mappings in neural systems devoted to perception, action and affective experience; a mechanism that can potentially realize the embodiment of language (Aryani, Conrad, et al., 2018; Meteyard, Stoppard, Snudden, Cappa, & Vigliocco, 2015; Perniss & Vigliocco, 2014; Vigliocco, Meteyard, Andrews, & Kousta, 2009; Vinson et al., 2015). In addition to behavioral studies supporting an iconic advantage for language learning (e.g., Imai et al., 2008) and language processing (e.g., Vinson et al., 2015), a growing number of neuroimaging research in the past few years aimed at revealing the neural mechanisms underlying such beneficial role of iconicity mostly by focusing on ideophones. In the present work, we aimed at extending this line of research to other types of words (e.g., regular nouns) by capitalizing on the role of affect in sound-to-meaning correspondences, which, as we will argue later, might be of crucial relevance to the research on iconicity in language.

1.1 | Neural evidence for a processing advantage of iconicity

Results of previous neuroimaging studies on iconicity indicate that iconic words profit from an additional processing network in the brain. In a series of fMRI-studies, Osaka and his colleagues provided some of the first neuroimaging studies on the topic (Osaka, 2009, 2011; Osaka & Osaka, 2009; Osaka, Osaka, Morishita, Kondo, & Fukuyama, 2004). Comparing Japanese mimetic words expressing laughter and pain with pseudowords, Osaka et al. showed that this type of iconic words activate the premotor brain areas associated with an actual laughter and pain, as well as, striatal reward area and cingulate cortex, respectively (Osaka et al., 2004; Osaka & Osaka, 2005). However, this series of experiments possesses a serious limitation with respect to comparing ideophones with pseudowords (Lockwood & Dingemanse, 2015): as embodiment theories assume that arbitrary words also activate relevant domain-specific sensorimotor areas (Hauk, Johnsrude, & Pulvermüller, 2004; Vigliocco et al., 2009; Zwaan, 2004), the potential advantage in processing of iconic words in these studies remained unclear. Kanero, Imai, Okuda, Okada, and Matsuda (2014) overcame this shortcoming by comparing onomatopoetic expressions that were related to motion and shape with arbitrary words from the same semantic domains. Results showed greater general activation, and a cluster of activation in the right posterior superior temporal sulcus (pSTS) presumably working as a hub for integration of multimodal (i.e., lexical and sublexical) information. This finding aligns with the results of previous work on onomatopoetic words showing a greater activation for bimodal information (i.e., onomatopoetic words imitating animal calls) in the left and right superior temporal sulcus (STS) than for unimodal information (i.e., either animal names, or animal calls; Hashimoto et al., 2006). By extending the word material to a multi-language stimulus set, Revill, Namy, DeFife, and Nygaard (2014) provided further support for the advantageous processing of iconic words and for the potential role of areas engaged in multimodal sensory integration beyond those involved in semantic processing. These results suggest the existence of more direct links between semantic information and sound information for iconic words with corresponding neural hubs as convergence zones for information integration. Iconic words might therefore be more immune to neurological damages that affect language-processing networks as, for instance, in aphasic patients. In fact, in a recent lesion study involving individuals with aphasia following left-hemisphere stroke (Meteyard et al., 2015), a consistent processing advantage was observed for onomatopoetic words in reading aloud and auditory lexical decision; two tasks that rely on sound–meaning mapping.

Overall, previous studies suggest that iconicity can facilitate language processing through activation of additional links between the sound of a word and modality-specific experiences (i.e., sensory, motor, and affective), as well as through integration of information from different modalities which may provide an opportunity for stronger embodiment of iconic signs (Aryani, Conrad, et al., 2018; Vigliocco et al., 2009). Nevertheless, unlike pioneering work on the facilitative effect of iconicity in sign language (Thompson, Vinson, Woll, & Vigliocco, 2012; Vinson et al., 2015) which laid the ground for the theoretical framework of such investigations, related research on spoken language—including behavioral studies—has so far mainly focused on onomatopoeia and ideophones including Japanese mimetic words (Dingemanse, Schuerman, Reinisch, Tufvesson, & Mitterer, 2016; Iwasaki, Vinson, & Vigliocco, 2007; Lockwood, Hagoort, & Dingemanse, 2016), or on cases typically not considered iconicity but rather statistical regularities in vocabulary (i.e., systematicity; Farmer, Christiansen, & Monaghan, 2006; Reilly, Westbury, Kean, & Peelle, 2012).

1.2 | The present study

Here, we focused on the iconic relationship between sound and meaning of words in the affective domain, termed affective iconicity (Aryani, 2018). As previously argued, we believe that extending the results of previous research on iconicity to the affective domain is important: Firstly, both affective communication (Darwin, 1871; Panksepp, 2010) and iconicity (Ramachandran & Hubbard, 2001; Reilly et al., 2012) have been considered crucial factors underlying language evolution. Thus, an iconic relationship between sound and meaning of words in today's language might be most evident in the affective domain. Secondly, affective dimensions of words, and in particular valence and arousal, are essential features defining a two-dimensional semantic space allowing for a very basic and potentially the most relevant distinction between different concepts (Osgood, 1952). In the present study, we therefore focused on affective iconic words and aimed at extending previous neural evidence on the facilitative role of iconicity in language processing to other classes of words than ideophones, and to the affective domain. Notably, the notion of "affective meaning" may not be shared by all theories on linguistic meaning. Our approach in this work is based on an embodied view of language which proposes that meaning is grounded in behavior and neural circuitry of the producer or the interpreter of linguistic signs, and that affective meaning is intertwined with other lexico-semantic aspects (Jacobs, Hofmann, & Kinder, 2016; Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012; Vigliocco et al., 2009, see Aryani, Conrad, et al., 2018, for more explanations).

For the present work, we built upon the results of two recent studies that capitalize on the affective domain showing an influence of the sound of words on the evaluation of affective meaning (Aryani, Conrad, et al., 2018), as well as a facilitative role of iconicity in online affective evaluation (Aryani & Jacobs, 2018). Based on the congruence versus incongruence of lexical (meaning) and sublexical (sound) arousal, words in the latter study were organized in two groups of iconic versus non-iconic, and presented in a two-alternative forced choice task regarding their lexical arousal (high vs. low). Results showed that iconic words were evaluated more quickly and more accurately, indicating a beneficial processing of iconic words compared to their non-iconic counterparts. This finding suggests that affective cues in the sound of a word (implicit or explicit) can be integrated with higher-order semantic processes facilitating the evaluation of affective content when sound and meaning aspects are congruent.

Using event-related fMRI in the present study, we aimed at exploring the neural mechanisms underlying the beneficial processing of affective iconic words. For this, we used a similar experimental design as in the aforementioned study, characterized by an orthogonal manipulation of *lexical* and *sublexical arousal* (representing meaning and sound, respectively; see Figure 1, and Materials and methods for details) and

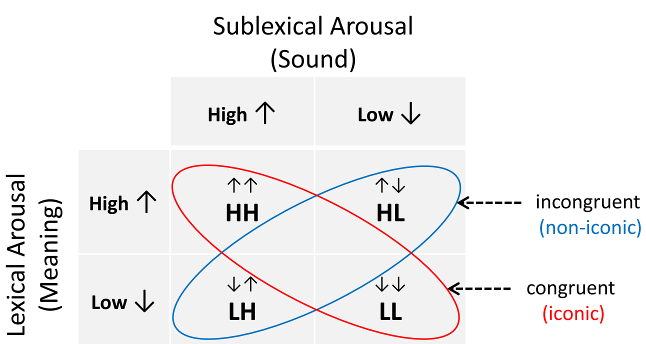

FIGURE 1 Word stimuli were organized in a 2 × 2 design: With each experimental factor (*lexical* and *sublexical arousal*, representing meaning and sound) manipulated in two distinctive groups consisting of extreme levels of arousal (high = exciting, and low = calming). The congruence versus incongruence of *lexical* (meaning) and *sublexical arousal* (sound) results in two groups of iconic versus non-iconic words regarding affective arousal [Color figure can be viewed at wileyonlinelibrary.com]

presented words in a passive listening task. We predicted a generally greater activation for iconic words than non-iconic words, particularly in brain areas associated with affective processing, and in convergence zones responsible for multimodal representation of emotional information from different sources: that is, acoustic information (related to *sublexical arousal*) and semantic information (related to *lexical arousal*).

Candidates for regions integrating emotional information from different domains have been found in previous work by focusing mostly on the integration of audiovisual cues, as well as verbal and nonverbal vocal cues. These extend from higher association areas such as the anterior and the posterior cingulate cortex (ACC and PCC), to prefrontal cortex (PFC), and (left) amygdala (Klasen, Chen, & Mathiak, 2012; Klasen, Kenworthy, Mathiak, Kircher, & Mathiak, 2011; Wittfoth et al., 2009). However, among these areas, PFC, and ACC responded more strongly to incongruent than congruent emotional information in accordance with their prominent role in the conflict network (Botvinick, Cohen, & Carter, 2004; Etkin, Egner, & Kalisch, 2011; Hofmann et al., 2008; Kerns et al., 2004). Therefore, we expected to observe activation in PCC and/or the left amygdala—as supramodal emotion integration networks—for iconic words due to the congruence between sublexical and lexical affective information.

To investigate the interaction between brain regions involved in iconic sound–meaning mappings, we also performed a functional connectivity analysis and adopted two independent seed regions representative for processing of sound (left superior temporal gyrus, STG) and meaning (inferior frontal gyrus, IFG) of words using functional masks extracted from the contrast between all words and the auditory baseline (see Materials and Methods). We hypothesized that iconicity significantly increases the coupling between these two seeds, on the one side, and the convergence zones integrating emotional information (PCC and/or the left amygdala), on the other side.

It is worth to point out that the use of a 2 × 2 design in this study enabled us to investigate (a) the main effect of lexical arousal, (b) the main effect of sublexical arousal, and (c) the effect of congruence between lexical and sublexical arousal (iconicity). Among these, the effect of lexical arousal has been well investigated in a number of previous studies (Etkin et al., 2011; Kuchinke et al., 2005; Lewis, Critchley, Rotshtein, & Dolan, 2007). Thus, the aim of the present study was to investigate the effect of sublexical arousal, which has been reported and discussed in a previous publication (Aryani, Hsu, & Jacobs, 2018), and the effect of congruence or iconicity, which is the focus of the present article.

2 | MATERIALS AND METHODS

2.1 | Stimuli

One hundred and twenty nouns (one to three syllables long) were selected for a 2 × 2 design (30 words for each condition) characterized by an orthogonal twofold manipulation of *lexical* and *sublexical arousal* (Figure 1). For *lexical arousal* we used ratings of words' affective meaning (minimum = 1: very low arousing, maximum = 5 very high arousing) from the normative database BAWL-R (Võ et al., 2009) which has been cross-validated in over hundred published studies regarding experiential, behavioral, and neurobiological levels of analysis (for review: Jacobs et al., 2015). *Sublexical arousal* was calculated based on features extracted from the acoustic representation of words applying the psychoacoustic model developed in a previous work (Aryani, Conrad, et al., 2018). The model is based on specific extracted acoustic features (e.g., pitch, formants, and intensity) of a list of almost 1,000 pseudowords and the ratings given on the affectivity of their sound. To convert the words in spoken form and to extract the corresponding acoustic features, a professional male actor was recruited, who was a native speaker of German. He was paid to participate. Words were spoken in a list-like manner to prevent affective prosody and were recorded in the "Leibniz-Zentrum Allgemeine Sprachwissenschaft" in Berlin in a professional sound recording booth using a "Sennheiser MKH20" microphone and "Ultra Gain MIC-2000" preamplifier. The audio signal was recorded using the DAT-recorder "TASCAM DA20MKII" with a sampling frequency of 48 kHz and 16 bits per sample. Spoken words were normalized to have the same loudness by matching their root-mean-square (RMS) power. The acoustic features of words were then extracted using the speech analysis software PRAAT (Boersma & Weenik, 1996) and used in the acoustic model to predict their *sublexical arousal*.

Words were divided into two distinctive conditions of "high" and "low" arousing for each of the factors *lexical arousal* ("High" > 3.25, "Low" <2.75) and *sublexical arousal* ("High" > 3, "Low" < 3), and carefully controlled for relevant psycholinguistic variables across all of four cells of experimental conditions. *Lexical arousal* (and *lexical valence*) was closely controlled for between the two cells of *sublexical arousal*, and vice versa (Table 1). This design resulted in four experimental conditions of HH, HL, LH, and LL (Figure 1). Example words for these conditions are: HH = Gemetzel /g ə m ɛ ts ə l/ (slaughter), Hitze /h ɪ ts ə/ (heat), HL = Elend /e: l ə n t/ (miserable), Lawine /l a: v ɪ n ə/ (avalanche), LH = Kreis /k r aɪ s/ (circle), Fassade (veneer) /f a s s a: d ə/, LL = Bohne /b o: n ə/ (bean), Sandale /z a n d a: l ə/ (sandal).

In order to create an acoustic baseline, we randomly selected 16 words from the word material (4 from each condition) and converted them to signal-correlated noise (SCN [Schroeder, 1968]). Along with our stimulus material (120 words +16 SCN), a total of 74 additional words (mostly emotionally neutral) were presented which were a part of another study, and were discarded from further analysis here.

2.2 | Participants

Twenty-nine right-handed German native speakers (17 women, mean age 25.2 years, range: 20–35 years) with no history of neurological or psychiatric illness or any hearing problems volunteered to participate in the study, receiving either 15 Euros or psychology course credit for their participation. The Ethical Committee of the Freie Universität Berlin had approved the investigation. Informed consent was obtained according to the Declaration of Helsinki.

2.3 | Procedure

Participants performed a passive listening task. Spoken words were presented via MRI-compatible headphones (RTC2K, from Resonance Technology Inc.) sufficiently shielded from scanner noise to ensure clear perceptibility. Participants were instructed to pay attention and to carefully listen to the words. Prior to word presentation, a fixation cross was presented on the center of the screen for between 1,500 and 6,500 ms, jittered in steps of 500 ms. Jittering durations and the stimulus presentation order over different experimental conditions (HH, HL, LH, LL, SCN, and Fillers), were optimized with Optseq2 to ensure a maximal signal-to-noise ratio (Greve, 2002). Accordingly, six different schedules (stimulus order combined with specific jittering durations) were generated and used across the subjects. The stimulus duration was 845 ms in average. After presentation of a stimulus the fixation cross disappeared. A total number of 10 trial words were

TABLE 1 Characteristics of word stimuli

	Word category								Inferential statistics	
	HH		HL		LH		LL			
Variable	M	SD	M	SD	M	SD	M	SD	F-test	Two sample t-tests
Lexical arousal	4.07	0.24	4.04	0.22	1.99	0.16	1.99	0.18	$F(3,116) = 983$, $p < .0001$	p (HH–HL) = .56, p (LL,LH) = .96
Lexical valence	−1.83	0.52	−1.83	0.51	0.22	0.36	0.18	0.37	$F(3,116) = 205$, $p < .0001$	p (HH–HL) = .93, p (LL,LH) = .72
Sublexical arousal	3.36	0.31	2.76	0.19	3.30	0.27	2.77	0.21	$F(3,116) = 50.5$, $p < .0001$	p (HH–LH) = .21, p (LL,HL) = .63
Word frequency	0.64	0.75	0.74	0.76	0.57	0.78	0.51	0.75	$F(3,116) = 0.47$, $p = .69$	.26 < ps < .78
Imageability rating	4.78	1.01	4.56	1.0	4.93	0.90	5.02	1.16	$F(3,116) = 1.11$, $p = .34$	.09 < ps < .74
# syllables	1.86	0.73	2.1	0.54	2.0	0.69	2.03	0.61	$F(3,116) = 0.68$, $p = .56$	.16 < ps < .84
# phonemes	5.3	1.36	5.23	1.10	5.13	0.89	4.93	1.20	$F(3,116) = 0.57$, $p = .63$	.22 < ps < .82
Duration (ms)	873	116	850	102	826	108	836	100	$F(3,116) = 1.06$, $p = .36$	.09 < ps < .72

Note: The first letter indicates the *lexical* and the second *sublexical arousal*.
Abbreviations: HH, high–high; HL, high–low; LH, low–high; LL, low–low.

presented prior to the experiment, which were excluded from the analysis. Words were split and presented in two runs between which the participants could take a break. A number of (120 + 16 + 76)/2 = 98 stimuli were presented in each run.

2.4 | fMRI data acquisition

Imaging data were collected on a Siemens Tim Trio 3 T MR scanner. Functional data used a T^*_2-weighted echo-planar sequence (slice thickness: 3 mm, no gap, 37 slices, repetition time [TR]: 2 s, echo time [TE]: 30 ms, flip angle: 70°, matrix: 64 × 64, field of view [FOV]: 192 mm, voxel size: 3.0 × 3.0 × 3.0 mm^3, 2 × 305 volumes, acquisition time: 2 × 610 s). At the beginning of the experimental session, magnitude and phase images for the field map were acquired: (slice thickness: 3 mm, no gap, 37 slices, TR: 488 ms, 2 TE: 4.92 and 7.38 ms, flip angle: 60°, matrix: 64 × 64, FOV: 192 mm, voxel size: 3.0 × 3.0 × 3.0 mm^3, acquisition time: 65 s). Individual high-resolution T1-weighted anatomical data (MPRAGE sequence) were also acquired (TR: 1.9, TE: 2.52, FOV: 256, matrix: 256 × 256, sagittal plane, slice thickness: 1 mm, 176 slices, resolution: 1.0 × 1.0 × 1.0 mm^3).

2.5 | Post-scan tests

At the end of the experiment, outside the scanner, an unannounced recognition test was performed to assess participants' involvement in the task and mnemonic effects of the experiment. After the recognition test, in order to check the reliability of our experimental manipulations participants were asked to evaluate the words presented in the scanner for their *lexical* and *sublexical arousal* in two separate rating studies.

2.5.1 | Unannounced recognition test

Participants were presented with the same 120 words used in the scanner (OLD) mixed with 120 new words (NEW) which were matched with OLD items for word frequency, number of letters, number of phonemes, number of syllables, and imageability rating, as well as valence and arousal (selected from the same range as used for OLD items). Participants were asked to rate how confident they were that the presented word was or was not part of the word list in the scanner (from certainly not presented in the scanner = 1 to certainly presented in the scanner = 5).

2.5.2 | Affective ratings

For rating of *lexical* and *sublexical arousal*, the same instruction was used as in the original rating study of the BAWL-R (Võ et al., 2009). For rating of *sublexical arousal*, participants were additionally instructed to only concentrate on the sound aspect of the words while trying to suppress their meaning (c.f. Aryani, Conrad, et al., 2018).

2.6 | fMRI preprocessing

The fMRI data were preprocessed and analyzed using the software package SPM12 (www.fil.ion.ucl.ac.uk/spm). Preprocessing consisted of slice-timing correction, realignment for motion correction, magnetic field inhomogeneity correction through the creation of a field map, and coregistration of the structural image onto the mean functional image. The structural image was segmented into gray matter, white matter, cerebrospinal fluid, bone, soft tissue, and air/background (Ashburner & Friston, 2005). A group anatomical template was created with DARTEL (Diffeomorphic Anatomical Registration using Exponentiated Lie algebra; Ashburner, 2007) toolbox from the segmented gray and white matter images. Transformation parameters for structural images were then applied to functional images to normalize them to the brain template of the Montreal Neurological Institute (MNI) in the original resolution of 3 × 3 × 3 mm. Normalized functional images were spatially smoothed with a Gaussian kernel of 6 mm full-width-at-half-maximum.

2.7 | fMRI analysis

2.7.1 | GLM analysis

Voxel-wise fixed effects contrast images made by subtraction analyses were performed at the single subject level and random effects analyses (Holmes & Friston, 1998) were conducted at the group level to create SPM contrast maps. On the single-subject level, each of the six conditions (HH, HL, LH, LL, SCN, and FILLERS) was convolved with the haemodynamic response function (HRF). Events were modeled as delta functions with zero duration. The beta images of each conditional regressor were then taken to the group level, where a full-factorial second level analysis with the factors *lexical arousal* and *sublexical arousal* was used. An unconstrained nondirectional 2 × 2 ANOVA whole brain analysis was performed with the factors *lexical arousal* (High, Low) and *sublexical arousal* (High, Low), to investigate the overall presence of main and interaction effects. For whole-brain fMRI analyses, according to the recent recommendations to avoid false positive results for cluster level correction (Eklund, Nichols, & Knutsson, 2016; Flandin & Friston, 2017), we used the cluster defining threshold (CDT) of p = .001, then applied cluster-level family-wise error (FWE) correction to p < .05 for the entire image volume. Small volume correction was also performed using region of interests (ROIs) defined based on the anatomical amygdala and PCC map using the WFU Pickatlas Tool.

The labels reported were taken from the "TD Labels" (Lancaster et al., 2000) or "aal" labels in the WFU Pickatlas Tool. The Brodmann areas (BA) were further checked with the Talairach Client using the nearest gray matter search after coordinate transformation with the WFU Pickatlas Tool.

2.7.2 | Functional connectivity analysis

To investigate regions showing significant functional connectivity with brain regions processing sound and meaning of words related to iconicity, generalized psychophysiophysiological interactions (gPPPI) (McLaren, Ries, Xu, & Johnson, 2012) were analyzed using the activations in regions representative for acoustic and semantic processing as seeds. For the sound aspect, we selected the left STG, and in particular auditory cortex, as a representative seed region for acoustic processing. For the meaning aspect, we focused on the left inferior frontal gyrus (IFG) due to its clear involvement in semantic processing (Binder, Desai, Graves, & Conant, 2009) and in appraisal of the words' affective meaning (Jacobs et al., 2016). These seed regions were defined based on the observed activations in the comparison of all words and acoustic baseline (see Results: Words > SCN). The left STG was extracted from a cluster of activation with an activation peak in [x y z] = [−59–14 3], and a cluster size of k = 437 voxels, and the left IFG from a cluster of activation with an activation peak in [x y z] = [−41 33–14], and a cluster size of k = 432 voxels. To analyze patterns of functional connectivity, we used the gPPI toolbox, which produces a design matrix with three columns of condition-related onsets with canonical HRF, BOLD signals deconvolved from the seed region, and PPPI regressors at the individual level. Thus, the GLM of this analysis included PPPI and condition regressors of the four experimental conditions (i.e., HH, HL, LH, and LL). We then performed group-level one-sample analysis of the contrast image of PPPI beta-maps between congruent and incongruent conditions. We checked results on the whole brain level and also performed SVC analysis with predefined ROIs of amygdala and PCC (see above).

3 | RESULTS

3.1 | Behavioral results

3.1.1 | Unannounced recognition test

Across all participants, we performed a linear mixed model (LMM) analysis predicting the recognition rate, with word category (OLD = “words used in the scanner” vs. NEW = “words not used in the scanner”) as fixed factor and words as well as participants as random factors. Results supported a performance above chance for recognizing OLD words, with a significantly higher score (M = 3.53) compared to NEW words (M = 2.54, t = −20.6, $p < .0001$). We next performed simple t-tests to compare the recognition rate between the levels of word category (OLD vs. NEW) separately for each participant. An effect of word category (OLD vs. NEW) on accuracy was observed for 27 participants out of 29 (t = 6.4 ± 3.2). These results indicate that the majority of participants had been attentive during the passive listening task. Two participants with a performance not higher than chance level (participant 4: t = 0.28, participant 19: t = 1.14) were consequently excluded from further analyses.

3.1.2 | Affective ratings

Lexical and *sublexical arousal* ratings used for stimulus construction were correlated with our post-scan data. For both, the correlation coefficients were very high: $r = .97$, $p < .0001$, (r_{min} among all participants = .73), and $r = .76$, $p < .0001$ (r_{min} among all participants = 0.49), respectively.

3.2 | Neuroimaging results

3.2.1 | GLM results

The comparison between all words contrasted with the auditory baseline condition (SCN, see Methods) revealed left-lateralized activations in core language areas, that is, the inferior frontal gyrus (IFG), middle and STG, and inferior parietal lobule (BA 40), suggesting that this experiment successfully tapped into the language processing system. Activity was also observed in bilateral parahippocampal gyrus, middle frontal gyrus and precentral gyrus, as well as the left superior frontal gyrus, the fusiform area, the right caudate, and superior parietal lobule.

Results of the two main effects of *lexical arousal* (Lex H > Lex L) and *sublexical arousal* (Sub H > Sub L) were reported and discussed in detail elsewhere (Aryani, Hsu, & Jacobs, 2018). In summary, the comparison Lex H > Lex L was associated with activation in brain regions involved in appraisal and general processing of affective stimuli, that is, dorsolateral and medial prefrontal cortex, PCC, the left inferior frontal gyrus (LIFG), and temporal pole. Brain regions associated with the comparison Sub H > Sub L were substantially similar to those involved in processing of other types of affective sounds, that is, superior temporal area, bilateral insula, and premotor cortex (Aryani, Hsu, & Jacobs, 2018), supporting a unifying view that suggests a core neural network underlying any type of affective sound processing (Frühholz, Trost, & Kotz, 2016). These results, together with those of the comparison Words > SCN, as well as of the behavioral studies strongly support the reliability of our experimental manipulations and show that this experiment successfully engaged participants in carefully listening to words.

For the present contrast between congruent and incongruent words (Congruent > Incongruent), following our hypothesis, we looked at the potential activation of amygdala and PCC using small volume correction (SVC) analysis. Results revealed no significant cluster of activation for PCC, but one for the left amygdala (FWE-corrected at the peak level, $p = 7 \times 10^{-5}$) with an activation peak in [x y z] = [−27–3–15], and a cluster size of k = 17 (Figure 2). At the whole brain level of analysis, a cluster of activation with an activation peak at the same voxel in the left amygdala ([x y z] = [−27–3–15]), and a cluster size of k = 58 survived the significant test (FWE-corrected at the cluster level, $p = .009$). None of the analyses revealed any cluster of activation for the reverse contrast (Incongruent > Congruent).

3.2.2 | Functional connectivity results (PPPI)

At the whole brain level, no cluster of activation survived the significant tests for the PPPI analysis between iconic and non-iconic words (Congruent > Incongruent). However, the SVC analysis based on the anatomical masks of amygdala and PCC revealed a significant cluster in the left amygdala (with an activation peak in [x y z] = [−28–4–13], $p = .002$, FWE-corr, Figure 3), but no activation cluster in PCC.

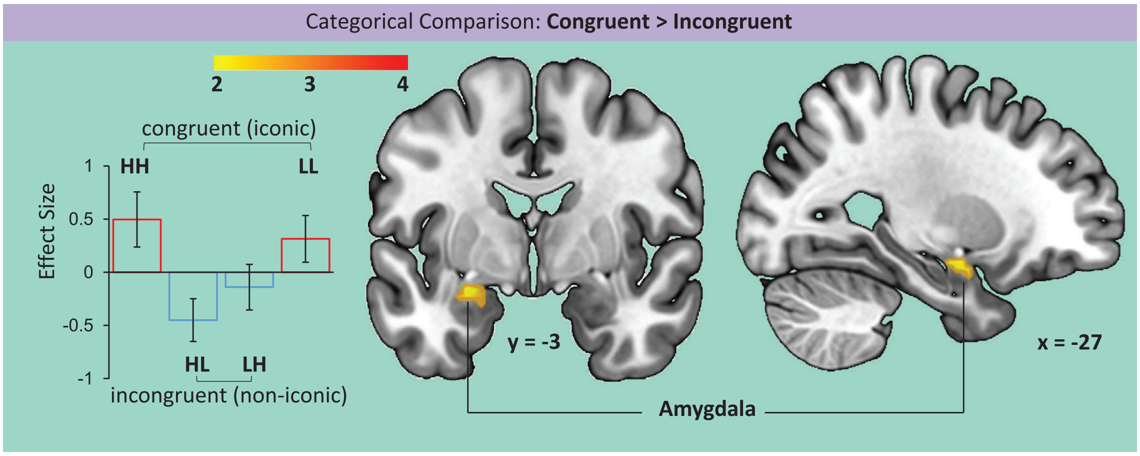

FIGURE 2 Iconic words as defined by the congruence between *lexical* and *sublexical arousal* elicited BOLD signals in the left amygdala ($p < .05$, FWE-corr). Left: Parameter estimates of the response in all conditions from the peak-activation voxel in [x y z] = [–27–3–15]. Pairwise comparisons showed increased activation in the same region for the contrast HH > HL, as well as LL > LH [Color figure can be viewed at wileyonlinelibrary.com]

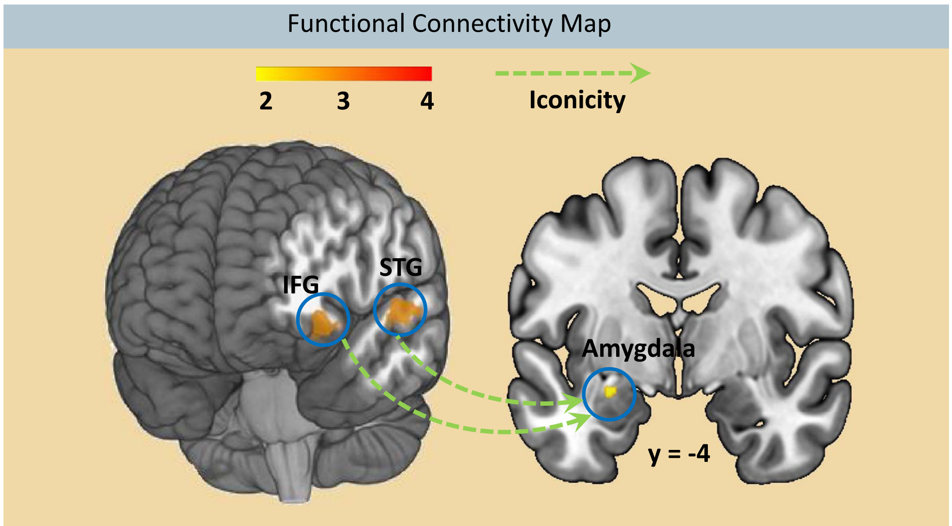

FIGURE 3 In the congruent condition (iconicity), the left amygdala showed significant functional connectivity with activation in two seed regions: The left superior temporal gyrus (STG) and the left inferior frontal gyrus (IFG) representing the processing of sound and meaning of words, respectively [Color figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

The present study examined to what extent iconic words—as defined by the congruence between affective sound and affective meaning—profit from an additional processing network that integrates the affective information in the sound and meaning of words. In line with our hypothesis, an interaction between affective information from two different sources (i.e., words' sublexical affective sound and lexico-semantic affective meaning) was observed as reflected in the left amygdala activity. Also, pairwise comparisons showed increased activation in the same region within two groups of lexically high and low arousing words (HH > HL, and LL > LH, see Figure 2). In addition, our functional connectivity analysis demonstrated that the observed activity in the left amygdala is modulated by activation in the left STG and the left IFG; two brain regions known for their prominent roles in sound and meaning processing.

The activation of the left amygdala in response to congruent emotional information from the sound and meaning of words is in line with its proposed role in supramodal emotion integration, and functioning as a general convergence zone (Klasen et al., 2011; Kreifelts, Ethofer, Huberle, Grodd, & Wildgruber, 2010; Müller, Cieslik, Turetsky, & Eickhoff, 2012; Schiller, Freeman, Mitchell, Uleman, & Phelps, 2009). The amygdala has reciprocal connections with association cortices in the superior and inferior temporal gyri (Aggleton, 1993) through which it can selectively modulate sensory responses depending on their emotional relevance. Thus, the amygdala appears to act as a neural gateway for binding the information from different modalities with each other, and also with brain region associated with emotional and motivational information. This view is in line with the modulatory role of amygdala in a wide array of networks and its functional importance in broader and more abstract dimensions of information processing (Jacobs et al., 2016; Pessoa & Adolphs, 2010).

Together with the previous finding regarding the main effects of lexical and sublexical arousal (Aryani, Hsu, & Jacobs, 2018), the overall results of our study show that this experiment successfully engaged the affective brain networks related to both levels of meaning (lexical) and sound (sublexical) processing. At the lexical level, our results reveal the involvement of dorsolateral and medial prefrontal cortex, as well as PCC, LIFG, and temporal pole known for their role in appraisal and general processing of affective stimuli (Etkin et al., 2011; Kuchinke et al., 2005; Kuhlmann, Hofmann, Briesemeister, & Jacobs, 2016; Lewis et al., 2007). At the sublexical level, the observed activation network is strongly similar to that related to other types of affective sounds, supporting the idea of a *unifying neural network* of affective sound processing (Frühholz et al., 2016). According to this view, all affective sounds consistently induce brain activity in a common core network, which consists of superior temporal cortex and amygdala, frontal and insular regions, and motor-related areas. Thus, our overall results not only provide a new neuroimaging evidence for the emotion potential of the sound of words (Aryani, Hsu, & Jacobs, 2018; Ullrich, Kotz, Schmidtke, Aryani, & Conrad, 2016), but they also enhance the chance of a general engagement of affective networks for processing of both aspects of words' sound and meaning, supporting the reliability of the subsequent results regarding the effect of congruence of these two aspects. Note that building a relation between mental states and neuroimaging activation patterns generally faces the problem of reverse inference (Poldrack, 2011) and correlational neuroimaging research, including the present work, cannot fully disentangle alternative interpretations. Therefore, more principled approaches are needed in future research that are, for instance, guided by computational models of the putative processes that underlie the psychological function (Vogt, 2018).

The results of our functional connectivity analysis share a substantial similarity to the results of previous work on neural mechanisms of affective prosody suggesting a network involving interactions between the STG and the IFG (Ethofer et al., 2011; Frühholz & Grandjean, 2013; Leitman et al., 2016) with the STG forming emotional representations of acoustic features (Aryani, Hsu, et al., 2018; Frühholz, Ceravolo, & Grandjean, 2011; Leitman et al., 2016; Wiethoff et al., 2008), and the IFG evaluating the meaning and the relevance of the sound (Leitman et al., 2016; Schirmer & Kotz, 2006). Importantly, these studies, in line with our results, suggest that the final appraisal of affective prosody takes place in the amygdala (Frühholz et al., 2016; Leitman et al., 2016) even though when emotional voices are presented outside the current focus of attention (Frühholz et al., 2011; Frühholz & Grandjean, 2013; Leitman et al., 2016).

The fact that we did not observe significant activation in the PCC, as expected in the introduction, might be due to the lack of socially relevant information in our stimuli. Previous studies on bimodal emotion integration mostly used human faces and voices both of which rely on social information potentially explaining the activation in the vPCC for emotionally congruent stimuli (Klasen et al., 2011; Schiller et al., 2009). However, unlike human faces and voices, the affective sound of words is based on basic acoustic features (Aryani, Conrad, et al., 2018) and is processed in substantially similar brain networks as other types of nonhuman affective sounds (Aryani, Hsu, et al., 2018).

Interestingly, unlike the results of multimodal integration of incongruent emotions, the inverse contrast for non-iconic words in our study (Incongruent > Congruent) did not elicit any significant cluster of activation. Although a neural effect of incongruent stimuli could intuitively be anticipated in brain regions associated with the conflict network, that is, PFC and ACC (Botvinick et al., 2004; Etkin et al., 2011; Hofmann et al., 2008; Kerns et al., 2004), the lack of significant activation for this contrast suggests that the human brain does not treat arbitrary relationships between sound and meaning as conflict. In fact, as the majority of words in the language are learned through conventional and per se arbitrary links, these results suggest that non-iconic words are chiefly processed in the core language regions, and even in the case of incongruence between sound and meaning no extra processing is devoted. Future research is needed to examine in more detail the neural substrates of incongruent words vs. a neutral baseline by using experimental designs with more distinct levels for *lexical* and *sublexical arousal* (e.g., high, medium, and low).

In addition to the activation increase for both iconic categories in the left amygdala, we found activation lower than the predefined perceptual baseline (Signal Correlated Noise, SCN; see Materials and Methods) in the same brain region for non-iconic categories (see Figure 2). "Deactivations" of this type are notoriously difficult to interpret (Frankenstein et al., 2003) particularly as the decrease in the BOLD signal for each of the categories is not absolute but relative to the acoustic baseline used in this experiment (i.e., Signal Correlated Noise, SCN; see Materials and Methods). Moreover, the pattern of response to an ongoing emotional stimulus in the amygdala has been shown to strongly depend on the level of oxygenation present in the foregoing condition and in the baseline (Whalen et al., 1998). Our suggestion is that the loss of signal strength observed for non-iconic conditions in the left amygdala is related to a potential signal increase in amygdala caused by our acoustic baseline. While SCN serves as a suitable acoustic baseline for investigating linguistic material (Davis & Johnsrude, 2003; Schroeder, 1968), its noisy characteristic and—in the present study—its unexpected presentation during each run might have an unpleasing effect on the participant similar to that of aversive sounds which are known to evoke brain responses in amygdala (Frühholz et al., 2015; Koelsch, 2014). Therefore, the negative values of the BOLD signal in amygdala for the non-iconic conditions (Figure 2) should be seen as a signal increase in the acoustic baseline relative to the non-iconic conditions.

A particular aspect of our study was extending the results of previous work that mainly focused on onomatopoeia and ideophones to another word class (i.e., regular nouns). A crucial issue of the use of onomatopoetic words in previous investigations on iconicity relates to the limitation of implicit testing of the sound effect. Usually possessing a special phonological construction, this type of words can possibly raise an undesirable awareness to the sound aspect of words when using in an experimental setting. Focusing on the affective domain, and in particular affective arousal, the present work provides quantitative measures for both sound and meaning of words, thereby enabling future research to apply the used method to any word class and to test sound–meaning correspondences across a wide range of words.

The present research is one of the first steps toward a better understanding of the neural mechanisms underlying the phenomenon of affective iconicity. It, however, poses a number of limitations that need to be addressed in future research. Firstly, iconic words in this study are defined as the congruence between lexical and sublexical arousal; the role of valence has therefore not been considered in this work. Given the inter-correlation between lexical arousal and valence, the actual contribution of each of the affective dimensions to the neural processes underpinning affective iconicity remains therefore largely unknown. Secondly, the fMRI data have been analyzed using conventional univariate methods, which focus on region-based activations. Thus, the use of more recent multivariate analysis methods such as representational similarity analysis (RSA; Kriegeskorte et al., 2008) can help move from localization of processes within regions to a better understanding of how these regions represent information. This approach is of particular relevance to this work as more than one feature in stimuli (i.e., valence and arousal) need to be mapped to neural activations. Lastly, future research can test the actual contribution of multimodal integration hubs to the processing of iconic words using transcranial magnetic stimulation (TMS) and the temporary-lesion paradigm. Such a stimulation would be expected to diminish the beneficial role of iconicity in semantic processing, that is, affective iconic words would be evaluated similarly to their non-iconic counterparts in a valence/arousal decision task (c.f., Aryani & Jacobs, 2018).

Showing a greater engagement of affective brain regions for (affective) iconic words, our finding can advance the understanding of affective and esthetic processes of literary reading (Aryani, Conrad, Jacobs, 2013; Aryani et al., 2016; Jacobs, 2015; Schrott & Jacobs, 2011). In line with its role in multimodal emotion integration, the left amygdala has been proposed to respond to metaphoric language, valence congruity, figurativeness, and harmony (Jacobs et al., 2016). Empirical support for this view comes from studies showing an enhanced left amygdala activation for metaphors (Citron & Goldberg, 2014) and metaphorical Noun-Noun-Compounds (Forgács et al., 2012) when compared to their literal counterparts. Also, results of a meta-analysis of 23 neuroimaging studies showed a left amygdala activation in response to a variety of figurative statements, and in particular metaphors (Bohrn, Altmann, & Jacobs, 2012). The meaning of metaphors is in general based on considerations of similarity between different aspects of target and source, and this is what iconicity in language is about. Turner and Lakoff (1989) defined iconicity as a "metaphorical image-mapping in which the structure of the meaning is understood in terms of the structure of the form of the language presenting the meaning." Such image-mapping, according to them, is enabled by image-schemas which are formed from our embodied experience. This view emphasizes the role of the left amygdala as a central hub critical for regulating the flow and integration of information from different experiences.

5 | CONCLUSION

The present data indicate that language users are sensitive to the interaction between sound and meaning aspect of words, and that the congruency of affective sound and affective meaning benefit from additional processing network. The corresponding neural mechanism potentially responsible for this sound–meaning interaction could be revealed in a brain area known for its role in multimodal emotion integration; that is, the left amygdala.

Some previous proposals restricted the role of iconicity to an earlier evolutionary stage of human language before the jump toward using a symbolic and arbitrary system, hence, considering iconicity a living "fossil" of proto-language. However, accumulating behavioral and neural evidence seriously challenges this claim and supports the existence and the neuropsychological reality of iconicity in today's language. Adding to this line of research, the present finding points to the indispensable role of iconicity in building a comprehensive knowledge about human language, and encourage future research to incorporate the underestimated iconic constituent of verbal symbols into linguistic theories, and to revisit the predominantly arbitrary character of language "awaiting due consideration in modern linguistic methodology" (Jakobson, 1965).

ACKNOWLEDGMENT

This research was partly funded by a grant to AA from the German Research Foundation (DFG) via the Cluster of Excellence "Languages of Emotion" at the Freie Universität Berlin. We would like to thank Markus Conrad for his crucial help in the conceptualization and the design of an earlier version of this project.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

DATA AVAILABILITY

The datasets generated and analyzed during the current study are available in the OSF repository at the following link: https://osf.io/u2xeq/?view_only=90b19a93b6074b5aa6f36db5133fd266

AUTHOR CONTRIBUTIONS

Study conception and design: A.A. and A.J., Acquisition of data: A.A., Analysis and interpretation of data: A.A., C.T.H., and A.J., Drafting of manuscript: A.A., Critical revision: A.A., C.T.H., and A.J.

ORCID

Arash Aryani https://orcid.org/0000-0002-0246-9422
Chun-Ting Hsu https://orcid.org/0000-0002-1829-6642

REFERENCES

Aggleton, J. P. (1993). The contribution of the amygdala to normal and abnormal emotional states. *Trends in Neurosciences*, *16*(8), 328–333.

Aryani, A. (2018). Affective Iconicity in Language and Poetry-A Neurocognitive Approach.

Aryani, A., Conrad, M., & Jacobs, A. M. (2013). Extracting salient sublexical units from written texts: "Emophon," a corpus-based approach to phonological iconicity. *Frontiers in Psychology*, *4*, 654, 1–15.

Aryani, A., Conrad, M., Schmidtke, D., & Jacobs, A. (2018). Why'piss' is ruder than'pee'? The role of sound in affective meaning making. *PLoS ONE*, *13*(6), e0198430.

Aryani, A., Hsu, C.-T., & Jacobs, A. M. (2018). The sound of words evokes affective brain responses. *Brain Sciences*, *8*(6), 94. https://doi.org/10.3390/brainsci8060094

Aryani, A., & Jacobs, A. M. (2018). Affective congruence between sound and meaning of words facilitates semantic decision. *Behavioral Sciences (Basel, Switzerland)*, *8*(6), 56.

Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A. M., & Conrad, M. (2016). Measuring the basic affective tone of poems via phonological saliency and iconicity. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(2), 191–204.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, *38*(1), 95–113. https://doi.org/10.1016/j.neuroimage.2007.07.007

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. https://doi.org/10.1016/j.neuroimage.2005.02.018

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*, 2767–2796. https://doi.org/10.1093/cercor/bhp055

Boersma, P., & Weenik, D. (1996). *PRAAT: A system for doing phonetics by computer. Report of the Institute of Phonetic Sciences of the university of Amsterdam*. Amsterdam: University of Amsterdam.

Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language-a quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, *50*, 2669–2683. https://doi.org/10.1016/j.neuropsychologia.2012.07.021

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Citron, F. M. M., & Goldberg, A. E. (2014). Metaphorical sentences are more emotionally engaging than their literal counterparts. *Journal of Cognitive Neuroscience*, *26*(11), 2585–2595.

Darwin, C. (1871). In C. Darwin (Ed.), *The descent of man, and selection in relation to sex*. London: Murray, 415. https://doi.org/10.1016/S0140-6736(01)31142-X

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *23*(8), 3423–3431. https://doi.org/23/8/3423 [pii].

De Saussure, F. (2011). *Course in general linguistics*. New York: Columbia University Press.

Deacon, T. (1997). *The symbolic species*. New York: Norton.

Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A Journal of General Linguistics*, *3*(1), 1–30. https://doi.org/10.5334/gjgl.444

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language*, *92*(2), e117–e133.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905.

Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., & Vuilleumier, P. (2011). Emotional voice areas: Anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. *Cerebral Cortex*, *22*(1), 191–200.

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, *15*(2), 85–93. https://doi.org/10.1016/j.tics.2010.11.004

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, *103*(32), 12203–12208. https://doi.org/10.1073/pnas.0602173103

Flandin, G., & Friston, K. J. (2017). Analysis of family-wise error rates in statistical parametric mapping using random field theory. *Human Brain Mapping*, *40*, 2052–2052. https://doi.org/10.1002/hbm.23839

Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M. J., Pléh, C., & Jacobs, A. M. (2012). Neural correlates of combinatorial semantic processing of literal and figurative noun noun compound words. *NeuroImage*, *63*, 1432–1442. https://doi.org/10.1016/j.neuroimage.2012.07.029

Frankenstein, U., Wennerberg, A., Richter, W., Bernstein, C., Morden, D., Rémy, F., & Mcintyre, M. (2003). Activation and deactivation in blood oxygenation level dependent functional magnetic resonance imaging. *Concepts in Magnetic Resonance*, *16A*(1), 63–70. https://doi.org/10.1002/cmr.a.10054

Frühholz, S., Ceravolo, L., & Grandjean, D. (2011). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex*, *22*(5), 1107–1117.

Frühholz, S., & Grandjean, D. (2013). Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neuroscience and Biobehavioral Reviews*, *37*(10), 2847–2855. https://doi.org/10.1016/j.neubiorev.2013.10.007

Frühholz, S., Hofstetter, C., Cristinzio, C., Saj, A., Seeck, M., Vuilleumier, P., & Grandjean, D. (2015). Asymmetrical effects of unilateral right or left amygdala damage on auditory cortical processing of vocal emotions. *Proceedings of the National Academy of Sciences*, *112*(5), 1583–1588. https://doi.org/10.1073/pnas.1411315112

Frühholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions-towards a unifying neural network perspective of affective sound processing. *Neuroscience and Biobehavioral Reviews*, *68*(June), 1–15. https://doi.org/10.1016/j.neubiorev.2016.05.002

Gasser, M. (2004). The origins of arbitrariness in language. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, *26*, 4–7.

Greve, D. N. (2002). Optseq2. Computer Software

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.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307.

Hockett, C. F. (1958). *A course in modern linguistics*. New Delhi: Oxford And Ibh Publishing Co.

Hofmann, M. J., Tamm, S., Braun, M. M., Dambacher, M., Hahne, A., & Jacobs, A. M. (2008). Conflict monitoring engages the mediofrontal cortex during nonword processing. *Neuroreport*, *19*(1), 25–29.

Holmes, A. P., & Friston, K. J. (1998). Generalisability, random E ects & population inference. *NeuroImage*, *7*, S754.

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

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007). What do English speakers know about Gera-Gera and yota-yota?: A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education around the Globe*, *17*(6), 53–78.

Jacobs, A. M. (2015). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*(April), 1–22. https://doi.org/10.3389/fnhum.2015.00186

Jacobs, A. M., Hofmann, M. J., & Kinder, A. (2016). On elementary affective decisions: To like or not to like, that is the question. *Frontiers in Psychology*, *7*(NOV), 1–17. https://doi.org/10.3389/fpsyg.2016.01836

Jacobs, A. M., Võ, M. L. H., Briesemeister, B. B., Conrad, M., Hofmann, M. J., Kuchinke, L., … Braun, M. (2015). 10 years of BAWLing into affective and aesthetic processes in reading: What are the echoes? *Frontiers in Psychology*, *6*(JUN), 1–15. https://doi.org/10.3389/fpsyg.2015.00714

Jakobson, R. (1965). Quest for the essence of language. *Diogenes*, *13*(51), 21–37. https://doi.org/10.1177/039219216501305103

Jakobson, R., & Waugh, L. R. (1979). *The sound shape of language*. Bloomington: Indiana University Press.

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.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*(5660), 1023–1026.

Klasen, M., Chen, Y. H., & Mathiak, K. (2012). Multisensory emotions: Perception, combination and underlying neural processes. *Reviews in the Neurosciences*, *23*(4), 381–392. https://doi.org/10.1515/revneuro-2012-0040

Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T. J., & Mathiak, K. (2011). Supramodal representation of emotions. *Journal of Neuroscience*, *31*(38), 13635–13643. https://doi.org/10.1523/JNEUROSCI.2833-11.2011

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, *15*(3), 170–180. https://doi.org/10.1038/nrn3666

Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., & Wildgruber, D. (2010). Association of trait emotional intelligence and individual fMRI-activation patterns during the perception of social signals from voice and face. *Human Brain Mapping*, *31*(7), 979–991. https://doi.org/10.1002/hbm.20913

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, *2*, 4.

Kuchinke, L., Jacobs, A. M., Grubich, C., Võ, M. L. H., Conrad, M., & Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: An fMRI study. *NeuroImage*, *28*(4), 1022–1032. https://doi.org/10.1016/j.neuroimage.2005.06.050

Kuhlmann, M., Hofmann, M. J., Briesemeister, B. B., & Jacobs, A. M. (2016). Mixing positive and negative valence: Affective-semantic integration of bivalent words. *Scientific Reports*, *6*, 1–7. https://doi.org/10.1038/srep30718

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., … Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, *10*(3), 120–131. https://doi.org/10.1002/1097-0193

Leitman, D. I., Edgar, C., Berman, J., Gamez, K., Fruhholz, S., & Roberts, T. P. L. (2016). Amygdala and insula contributions to dorsal-ventral pathway integration in the prosodic neural network, (215), 1–12. Retrieved from http://arxiv.org/abs/1611.01643

Lewis, P. A., Critchley, H. D., Rotshtein, P., & Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral Cortex*, *17*(3), 742–748. https://doi.org/10.1093/cercor/bhk024

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: A review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in Psychology*, *6*, 1246.

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How iconicity helps people learn new words: Neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra: Psychology*, *2*(1), 7. http://doi.org/10.1525/collabra.42

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*, *61*(4), 1277–1286.

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. https://doi.org/10.1016/j.cortex.2010.11.002

Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. *Neuropsychologia*, *76*, 264–275. https://doi.org/10.1016/j.neuropsychologia.2015.01.042

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), 325–347. https://doi.org/10.1037/a0022924

Müller, V. I., Cieslik, E. C., Turetsky, B. I., & Eickhoff, S. B. (2012). Crossmodal interactions in audiovisual emotion processing. *NeuroImage*, *60*(1), 553–561. https://doi.org/10.1016/j.neuroimage.2011.12.007

Osaka, N. (2009). Walk-related mimic word activates the extrastriate visual cortex in the human brain: An fMRI study. *Behavioural Brain Research*, *198*(1), 186–189. https://doi.org/10.1016/j.bbr.2008.10.042

Osaka, N. (2011). Ideomotor response and the neural representation of implied crying in the human brain: An fMRI study using onomatopoeia. *Japanese Psychological Research*, *53*(4), 372–378. https://doi.org/10.1111/j.1468-5884.2011.00489.x

Osaka, N., & Osaka, M. (2005). Striatal reward areas activated by implicit laughter induced by mimic words in humans: A functional magnetic resonance imaging study. *Neuroreport*, *16*(15), 1621–1624.

Osaka, N., & Osaka, M. (2009). Gaze-related mimic word activates the frontal eye field and related network in the human brain: An fMRI study. *Neuroscience Letters*, *461*(2), 65–68. https://doi.org/10.1016/j.neulet.2009.06.023

Osaka, N., Osaka, M., Morishita, M., Kondo, H., & Fukuyama, H. (2004). A word expressing affective pain activates the anterior cingulate cortex in the human brain: An fMRI study. *Behavioural Brain Research*, *153*, 123–127.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, *49*(3), 197–237.

Panksepp, J. (2010). Emotional causes and consequences of social-affective vocalization. In *Handbook of behavioral neuroscience* (Vol. 19, pp. 201–208). Amsterdam: Elsevier.

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*(DEC), 1–15. https://doi.org/10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *369*(1654), 20140179–20140179. https://doi.org/10.1098/rstb.2014.0179

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a'low road'to'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773–782.

Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron*, *72*(5), 692–697.

Ramachandran, V. S., & Hubbard, E. H. (2001). Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3–34. https://doi.org/10.1111/1468-0068.00363

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), 1–15. https://doi.org/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. https://doi.org/10.1016/j.bandl.2013.11.002

Schiller, D., Freeman, J. B., Mitchell, J. P., Uleman, J. S., & Phelps, E. A. (2009). A neural mechanism of first impressions. *Nature Neuroscience*, *12*(4), 508–514. https://doi.org/10.1038/nn.2278

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Schmidtke, D. S., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in Psychology*, *5*, 1–6. https://doi.org/10.3389/fpsyg.2014.00080

Schroeder, M. (1968). Reference signal for signal quality studies. *The Journal of the Acoustical Society of America*, *44*(6), 1735–1736.

Schrott, R., & Jacobs, A. M. (2011). *Gehirn und Gedicht: Wie wir unsere Wirklichkeiten konstruieren*. Munich: Hanser.

Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British sign language. *Psychological Science*, *23*(12), 1443–1448. https://doi.org/10.1177/0956797612459763

Turner, M., & Lakoff, G. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.

Ullrich, S., Kotz, S. A., Schmidtke, D. S., Aryani, A., & Conrad, M. (2016). Phonological iconicity electrifies: An ERP study on affective sound-to-meaning correspondences in German. *Frontiers in Psychology*, *7*, 1–17. https://doi.org/10.3389/fpsyg.2016.01200

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*(2), 219–247. https://doi.org/10.1515/LANGCOG.2009.011

Vinson, D., Thompson, R. L., Skinner, R., & Vigliocco, G. (2015). A faster path between meaning and form? Iconicity facilitates sign recognition and production in British sign language. *Journal of Memory and Language*, *82*, 56–85. https://doi.org/10.1016/j.jml.2015.03.002

Võ, M. L. H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior Research Methods*, *41*(2), 534–538. https://doi.org/10.3758/BRM.41.2.534

Vogt, N. (2018). Machine learning in neuroscience. *Nature Methods*, *15*(1), 33.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, *18*(1), 411–418. https://doi.org/9412517

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody-influence of acoustic parameters and arousal. *NeuroImage*, *39*(2), 885–893. https://doi.org/10.1016/j.neuroimage.2007.09.028

Wittfoth, M., Schröder, C., Schardt, D. M., Dengler, R., Heinze, H.-J., & Kotz, S. A. (2009). On emotional conflict: Interference resolution of happy and angry prosody reveals valence-specific effects. *Cerebral Cortex*, *20*(2), 383–392.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation*, *44*, 35–62.

How to cite this article: Aryani A, Hsu C-T, Jacobs AM. Affective iconic words benefit from additional sound–meaning integration in the left amygdala. *Hum Brain Mapp*. 2019;40:5289–5300. https://doi.org/10.1002/hbm.24772