This article was downloaded by: [University of Wisconsin - Madison]
On: 03 January 2013, At: 08:53
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Cognition & Emotion
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pcem20

Inherent emotional quality of human speech sounds

Blake Myers-Schulz [a], Maia Pujara [a], Richard C. Wolf [a] & Michael Koenigs [a]
[a] Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, USA
Version of record first published: 03 Jan 2013.

To cite this article: Blake Myers-Schulz , Maia Pujara , Richard C. Wolf & Michael Koenigs (2013): Inherent emotional quality of human speech sounds, Cognition & Emotion, DOI:10.1080/02699931.2012.754739

To link to this article: http://dx.doi.org/10.1080/02699931.2012.754739

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

COGNITION AND EMOTION, 2013
http://dx.doi.org/10.1080/02699931.2012.754739

Routledge
Taylor & Francis Group

BRIEF REPORT

Inherent emotional quality of human speech sounds

Blake Myers-Schulz, Maia Pujara, Richard C. Wolf, and Michael Koenigs

Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, USA

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

During much of the past century, it was widely believed that phonemes—the human speech sounds that constitute words—have no inherent semantic meaning, and that the relationship between a combination of phonemes (a word) and its referent is simply arbitrary. Although recent work has challenged this picture by revealing psychological associations between certain phonemes and particular semantic contents, the precise mechanisms underlying these associations have not been fully elucidated. Here we provide novel evidence that certain phonemes have an inherent, non-arbitrary emotional quality. Moreover, we show that the perceived emotional valence of certain phoneme combinations depends on a specific acoustic feature—namely, the dynamic shift within the phonemes' first two frequency components. These data suggest a phoneme-relevant acoustic property influencing the communication of emotion in humans, and provide further evidence against previously held assumptions regarding the structure of human language. This finding has potential applications for a variety of social, educational, clinical, and marketing contexts.

Keywords: Language; Emotion; Speech; Phoneme.

Human vocalisations can convey emotion through the semantic content of particular phoneme combinations (words), as well as through the prosodic features of a vocalisation (accentuation, intonation, rhythm). Throughout much of the past century, it was generally accepted that a string of phonemes (word) constitutes a linguistic sign via an arbitrary relationship between the phonemes and the referent, whereas prosody is thought to have an inherent, non-arbitrary emotional quality that is wholly independent from the semantic content (or lack thereof) in the vocalisation. To illustrate, the string of phonemes used to identify a particular object may vary widely between languages, while a particular prosody can be universally understood to convey a particular emotion. Recent work, however, has begun to challenge this picture. In this study, we demonstrate for the first time that certain strings of English phonemes have an inherent, non-arbitrary emotional valence that can be predicted on the basis of dynamic changes in acoustic features.

Correspondence should be addressed to: Michael Koenigs, Department of Psychiatry, University of Wisconsin-Madison, 6001 Research Park Blvd., Madison, WI 53719, USA. E-mail: mrkoenigs@wisc.edu

© 2013 Taylor & Francis

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

Although previous studies have identified a number of psychological associations between certain phonemes and particular semantic or perceptual characteristics, these studies differ from ours by employing a classification scheme based either on articulation features (i.e., configurations of the vocal tract during the production of phonemes) or on the relative position and dispersion of phonemes' acoustic frequency components (also known as their formant frequencies). In contrast, our study employs a classification scheme that is based on the dynamic shifts in the phonemes' formant frequencies.

One of the most consistent findings is that nonsense words consisting of front phonemes—i.e., phonemes articulated toward the front of the mouth—are perceived as smaller, faster, lighter, and more pleasant than nonsense words consisting of back phonemes (Folkins & Lenrow, 1966; Miron, 1961; Newman, 1933; Sapir, 1929; Shrum & Lowrey, 2007; Thompson & Estes, 2011). Such differences have also been found when comparing voiced and voiceless phonemes (Folkins & Lenrow, 1966; Klink, 2000; Newman, 1933; Thompson & Estes, 2011), rounded and unrounded phonemes (Kohler, 1929; Ramachandran & Hubbard, 2001), plosives and continuants (Westbury, 2005), and plosives and fricatives (Klink, 2000). Moreover, these results are not unique to English-speaking subjects, nor to adults (Davis, 1961; Hinton, Nichols, & Ohala, 1994; Maurer, Pathman, & Mondloch, 2006; though see I. K. Taylor & Taylor, 1965).

These findings raise the question of whether such effects result from some idiosyncratic physical feature of articulation (e.g., proprioceptive awareness of vocal-tract position, or the visual perception of associated facial expressions), or whether these effects are instead strictly related to the listener's auditory experience of acoustic features. The previous articulation-based findings do not resolve this issue; the main confounding factor is that differences in articulation features correspond systematically to differences in acoustic features. For example, as one moves from front to back vowels, the position of the second formant frequency (i.e., F2), as well as the dispersion between the first and second formants (i.e., F1 and F2), decreases at each step (Delattre, Liberman, & Cooper, 1955). Thus, the articulation-based finding—that nonsense words consisting of front phonemes are perceived as smaller, faster, lighter, and more pleasant than nonsense words consisting of back phonemes—could alternatively be described as an acoustic-based finding that nonsense words with a lower F2-position (or a smaller dispersion between F1 and F2) are perceived as smaller, faster, lighter, and more pleasant than nonsense words with higher F2-positions (or a greater dispersion between F1 and F2). To avoid such confounds, we employ a classification scheme that (i) is based on the acoustic features of phonemes and (ii) does not correspond to any single classification scheme regarding articulation features (e.g., the front–back continuum).

This focus on acoustic features also has the advantage of linking our study to previous proposals that have emphasised a key role for acoustic features. Work on animal vocalisation, for instance, suggests an important role for acoustic features in some of the previous findings on the semantic qualities of nonsense words. Several studies reveal a positive correlation between vocal-tract length and body size, in both human and non-human animals (Fitch, 1997, 1999; Riede & Fitch, 1999). Moreover, animals with longer vocal tracts (and thus larger body sizes) tend to produce vocalisations with lower formant frequencies and a smaller dispersion between formants (Charlton et al., 2011; Charlton, Reby, & McComb, 2008; Fitch, 2000; Fitch & Reby, 2001). A variety of species—including human, red deer, domestic dog, and koala—rely on these acoustic features to approximate the body size of other organisms (Charlton, Ellis, Larkin, & Fitch, 2012; Reby et al., 2005; Smith, Patterson, Turner, Kawahara, & Irino, 2005; A. M. Taylor, Reby, & McComb, 2010). Thus, a parallel can be drawn between people's tendency to associate certain nonsense words with particular semantic properties and the tendency of animals (human and nonhuman) to use formant features as cues for body size. Specifically, just as front

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

vowels (i.e., vowels characterised by a higher F2-position and a greater dispersion between F1 and F2) are perceived as smaller and lighter than back vowels (i.e., vowels characterised by a lower F2-position and a smaller dispersion between F1 and F2), a similar point can be made about certain animal vocalisations: namely, that animal vocalisations characterised by higher formant positions and a greater formant dispersion are perceived as originating from a smaller organism than vocalisations characterised by lower formant positions and a smaller formant dispersion. Accordingly, some authors have suggested that people's tendency to perceive front vowels as smaller and lighter than back vowels may be evolutionarily rooted in this use of formant position and dispersion as a cue for body size (see Shrum & Lowrey, 2007, for a review). In light of such proposals, the question arises as to whether there might be other acoustic features that play a similar type of dual role as that played by formant position and dispersion; i.e., serving as both the characteristic feature used to identify particular human phonemes (in the way that formant position and dispersion are used to identify vowel sounds) and the characteristic feature used to identify certain semantic properties in animal vocalisations (in the way that formant position and dispersion are also used as cues for body size). If an additional feature along these lines were found, it would raise the possibility that the semantic properties associated with this feature's expression in animal vocalisations might also be associated with this feature's expression in nonsense words (e.g., just as the semantic property of smallness is associated both with animal vocalisations characterised by high formant position and dispersion and with nonsense words characterised by high formant position and dispersion). One candidate for an additional feature along these lines is the transition patterns of formant frequencies (also known as formant shifts).

Our study's hypothesis is predicated on three observations. First, numerous non-human animal species lower their vocal tracts (thereby lowering the frequencies of their vocalisations) in order to appear larger and more threatening to antagonists or competitors (Fitch, 1999; Reby et al., 2005). Second, if the vocal tract is dynamically lengthened and pitch dynamically lowered during the presentation of a given vowel sound, humans perceive the vowel as originating from an angrier organism, whereas if the vocal tract is dynamically shortened and pitch dynamically raised during the presentation of the vowel sound, humans perceive the vowel as originating from a happier organism (Chuenwattanapranithi, Xu, Thipakorn, & Maneewongvatana, 2008). Third, in humans, the categorical perception of phonetic sounds is mediated not only by the relative position and dispersion of the first and second formant frequencies (F1 and F2), but also by the transition patterns of F1 and F2 (Delattre et al., 1955; Stevens, Blumstein, Glicksman, Burton, & Kurowski, 1992). Thus, we predicted that human vocalisations exhibiting a downward shift in F1/F2 formants (perhaps evolutionarily rooted in antagonistic/competitive behaviour) would be associated with negative emotion, whereas vocalisations exhibiting an upward shift in F1/F2 formants (perhaps evolutionarily rooted in conciliatory/submissive behaviour) would be associated with positive emotion. This prediction differs from the findings of Chuenwattanapranithi et al. (2008) in that Chuenwattanapranithi et al. examined changes in acoustic properties (such as pitch) that had no effect on subjects' differentiation between phonemes. In contrast, our study focused on certain acoustic properties (i.e., F1/F2 formant shifts) that play a key role in the differentiation between phonemes. This difference in focus allowed us to explore the novel hypothesis that certain phoneme-relevant acoustic features exhibit a non-arbitrary emotional quality.

METHODS

To test this prediction, we adapted a two-alternative forced-choice test in which subjects were instructed to match strings of phonemes (comprising nonsense words) with pictures (Kohler, 1929). The non-words were constructed

so as to exhibit either an overall upward or downward shift in F1/F2 frequencies (Figure 1, Table 1), and the pictures were selected on the basis of eliciting positive or negative emotion (Bradley, Codispoti, Cuthbert, & Lang, 2001; see also Table S1). Non-word pairs were matched for other acoustic and speech features (volume, intonation, accentuation, plosives, fricatives, nasals, vowels, number of syllables, number of phonemes, number of graphemes; Table S2), and pictures were matched for arousal. We ran two versions of the task with two different sets of healthy adult participants. In one version the word pairs were presented visually ($n = 32$ adult subjects, 15 males, M_{age} 35.5 ± 15.2 years), and in the other version the word pairs were presented auditorily ($n = 20$ adult subjects, 10 males, M_{age} 53.3 ± 11.7 years).

The visual test. This consisted of 35 trials (20 experimental, 15 control), corresponding to the 35 non-word pairs (Table 1) as well as the 35 image pairs (Table S1). For each of the 20 experimental trials, there was a distinct pair of non-words with opposing F1/F2 frequency shifts (one with upward shifts, one with downward shifts) and a distinct pair of pictures with opposing valences (one positive, one negative). For each of the 15 control trials, there was a distinct pair of non-words with opposing F1/F2 frequency shifts (one upward, one downward) and a distinct pair of pictures with matched valences (five of the pairs were neutral/neutral, five were negative/negative, and five were positive/positive). The 15 control trials (which were randomly interspersed among the experimental trails) had image pairs with matched (rather than opposing) valences in order to mask the semantic feature under scrutiny (i.e., valence). Thus, for control trials, the non-words within each pair were matched to positive, negative, and neutral images equally often.

The auditory test. This consisted of 22 trials. For each trial, there was a distinct pair of non-words with opposing F1/F2 frequency shifts (one upward, one downward) and an individual picture (11 trials had a positively valenced picture, and 11 had a negatively valenced picture). The non-words were pronounced by an automated computer voice program (AT&T Natural Voices). For each of the non-word pairs used in the auditory task, spectrograms were obtained with Praat software (Version 5.1.20) to ensure that the non-words in each pair differed with respect to direction of F1/F2 shift. The 22 non-word pairs used in the auditory task are italicised in Table 1. The 22 pictures used in the auditory task are italicised in Table S1. Given that the auditory task used individual pictures (rather than image pairs), there was no need to include control trials that matched image pairs for valence (as was done in the visual task).

In both the visual and auditory tests, the non-word pairs remained constant across all subjects. To illustrate, the two non-words in the first pair from Table 1—"babopu" and "tatoku"—were always paired together. However, the place at which each non-word appeared on the screen (top left or bottom right, for the visual task; bottom left or bottom right, for the auditory task) was randomised for each subject, as was the image

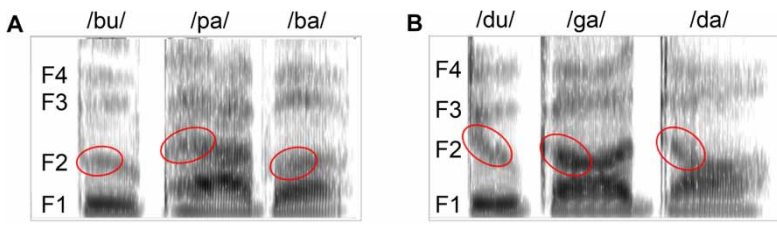

Figure 1. *Spectrograms illustrating the first four formants (F1–F4) of the example nonsense words "bupaba" (A) and "dugada" (B) as obtained with Praat software (Version 5.1.20). The differentiation between these non-words is largely mediated by the F2 transitions from consonants to vowels (outlined in red), which move upward for "bupaba" and downward for "dugada".*

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

Table 1. *Non-word pairs used in the study*

	Upward F1/F2 shifts (Predicted positive)	*Downward F1/F2 shifts (Predicted negative)*		*Upward F1/F2 shifts (Predicted positive)*	*Downward F1/F2 shifts (Predicted negative)*
1	*babopu*	*tatoku*	19	*mesabo*	*nezago*
2	*pabu*	*dagu*	20	*bepaso*	*tekazo*
3	*bopo*	*koto*	21	*bapafo*	*katavo*
4	*mobapo*	*nodago*	22	*mipaba*	*nidaga*
5	*bibapo*	*didago*	23	besa	teza
6	*masa*	*naza*	24	pafobu	tavoku
7	*pabapo*	*katako*	25	pasa	kaza
8	*mabapu*	*nadaku*	26	besapa	dezaga
9	*bafaso*	*gavazo*	27	asofi	azovi
10	*mabo*	*nago*	28	bofoba	kovoka
11	*bafapo*	*davako*	29	asaba	azaga
12	*bepabo*	*degado*	30	bobapu	kokatu
13	*pasobi*	*tazogi*	31	befosa	kevoza
14	*bopa*	*koga*	32	mabapu	nataku
15	*bobipa*	*dodiga*	33	ousa	ouza
16	*pafabi*	*davagi*	34	posabu	kozagu
17	*bupaba*	*dugada*	35	bapibu	gadigu
18	*bobipu*	*dodigu*			

Notes: Each of the 35 non-word pairs was used in the visual task. The subset of 22 italicised non-word pairs was used in the auditory task. Non-word pairs had opposing F1/F2 shifts, but were otherwise matched for acoustic and speech features (volume, intonation, accentuation, plosives, fricatives, nasals, vowels, number of syllables, number of phonemes, number of graphemes; for more detail see methods and Table S2).

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

pair (or the individual image, for the auditory task) that appeared with the non-words and the particular trial in which the non-word pair was presented.

Non-word pairs were constructed on the basis of the following general points:

1. The differentiation between bilabial plosives (/b/, /p/), alveolar plosives (/d/, /t/), and velar plosives (/g/, /k/) is largely mediated by the F2 transitions from the plosive to the subsequent vowel sound. In short, the F2 transitions of bilabial plosives tend to move upward, while the F2 transitions of alveolar and velar plosives tend to move downward (Blumstein & Stevens, 1979; Delattre et al., 1955; Stevens & Blumstein, 1978; Sussman, Bessell, Dalston, & Majors, 1997).
2. The differentiation between the nasal phonemes, /m/ and /n/, is largely mediated by the F2 transitions from the nasal to the subsequent vowel sound. Furthermore, the F2 transitions of /m/ and /n/ closely resemble the F2 transitions of /b/ and /d/, respectively (Delattre et al., 1955; Liberman, Delattre, Cooper, & Gerstman, 1954; Repp & Svastikula, 1988).
3. The differentiation between voiced intervocalic fricatives (particularly, /z/ and /v/) and their voiceless counterparts (/s/ and /f/, respectively) is largely mediated by the F1 transitions between the fricative and the adjacent vowels. Moreover, for the voiced fricatives /z/ and /v/, the downward F1 transition from the preceding vowel to the fricative tends to be greater than the upward F1 transition from the fricative to the subsequent vowel. For the voiceless fricatives /s/ and /f/, however, the downward F1 transition from the preceding vowel to the fricative tends to be less than the upward F1 transition from the fricative to the subsequent vowel. Furthermore, the downward F1 transition of fricatives

tends to be more salient than the upward F1 transition; and the voiced fricatives (/z/ and /v/) exhibit a greater downward F1 transition than their voiceless counterparts (/s/ and /f/, respectively; Stevens et al., 1992).

In accordance with (1) to (3), all non-words were paired so that: (i) bilabial plosives were matched with alveolar or velars plosives; (ii) the nasal phoneme /m/ was matched with /n/; (iii) the voiced fricatives (/z/ and /v/) were matched with their voiceless counterparts (/s/ and /f/, respectively); and (iv) all fricatives were situated between two vowels (hence, making them intervocalic fricatives). Also, given that the extent (and in some cases, the direction) of F1/F2 transitions can vary depending on the adjacent vowel sound (Delattre et al., 1955), and given that certain changes in vowel sounds can create changes in the perceived meaning of a nonsense word (Folkins & Lenrow, 1966; Miron, 1961; Newman, 1933; Sapir, 1929), all non-word pairs were matched for vowel sounds (as seen in Table 1).

RESULTS

As predicted, subjects reliably paired the downward F1/F2 shift non-words with the negative images and the upward F1/F2 shift non-words with the positive images. In the visual task ($n = 32$ adult subjects, 15 males, M_{age} 35.5 ± 15.2 years), subjects made the predicted choice on 80% of trials (chance performance is 50%). The proportion of individuals selecting a majority of predicted responses (i.e., on more than 10 out of the 20 trials) was significantly greater than expected by chance (Yates' $\chi^2 = 23.8$; $p = .000001$; Figure 2).

To ensure that this effect did not result merely from the words' visual properties, we ran an auditory version of the test with a different set of subjects ($n = 20$ adult subjects, 10 males, M_{age} 53.3 ± 11.7 years), and again found the predicted pattern. In particular, subjects made the predicted choice on 65.5% of trials (chance performance is 50%). The proportion of individuals selecting a majority of predicted responses (i.e., on more than 11 out of the 22 trials) was again significantly greater than expected by chance (Yates' $\chi^2 = 4.7$; $p = .03$; Figure 3).

DISCUSSION

These results suggest that certain strings of English phonemes have a non-arbitrary emotional quality, and, moreover, that the emotional quality can be predicted on the basis of specific acoustic features. This study moves beyond previous research by employing a classification scheme that is based solely on the acoustic features of phonemes, such that the scheme cannot be equivalently redescribed in terms of articulation features (e.g., the front–back continuum). Moreover, the present results differ from previous findings relating phoneme-*ir*relevant

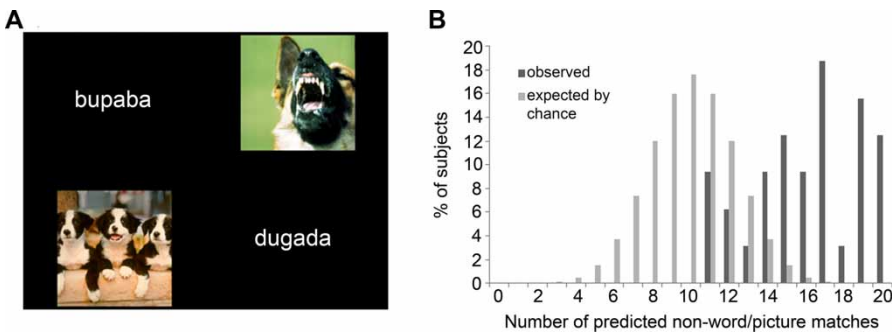

Figure 2. *Visual task. (A) Example of a trial from the visual task. Subjects pressed a vertical arrow button to match the non-words/pictures vertically, or a horizontal arrow button to match the non-words/pictures horizontally. (B) Data from the visual task.*

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

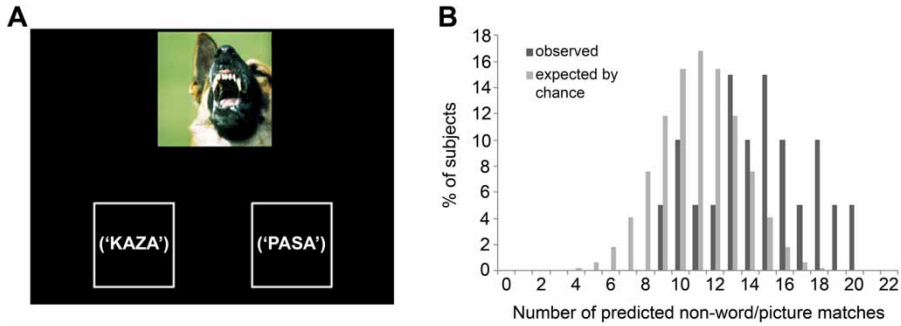

Figure 3. *Auditory task. (A) Example of a trial from the auditory task. Subjects heard two non-words (the non-words were not actually displayed visually during the task), then pressed a button (left or right) to select the first or second word, respectively. (B) Data from the auditory task.*

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

properties (such as pitch) to emotional valence (Chuenwattanapranithi et al., 2008); and also from findings relating certain acoustic features (e.g., amplitude) to emotional arousal, irrespective of valence (Bachorowski & Owren, 1995, 1996). This study is thus the first to identify phoneme-relevant formant shifts as a critical determinant of emotional valence in human speech.

The present results also link nicely to previous proposals that emphasise an important role for acoustic features. As noted, some authors have suggested that people's tendency to perceive front vowels as smaller and lighter than back vowels may be explained in terms of the parallel finding that animal vocalisations characterised by higher formant positions and a greater formant dispersion are perceived as originating from a smaller organism than vocalisations characterised by lower formant positions and a smaller formant dispersion. In a similar fashion, the present findings that strings of phonemes characterised by downward F1/F2 shifts are perceived as more aversive than those characterised by upward F1/F2 shifts may be explained in terms of parallel data suggesting that animal vocalisations characterised by a dynamic lowering of the vocal tract are perceived as angrier than vocalisations characterised by a dynamic raising of the vocal tract.

One limitation of the present study is that all subjects were English-speaking adults. Thus, future work would do well to examine whether the effect also holds in non-English speaking subjects and in younger children; especially since some of the key previous findings on the semantic properties of nonsense words have been found to extend to both non-English speaking subjects and toddlers (Davis, 1961; Hinton et al., 1994; Maurer et al., 2006).

The data presented here effectively outline a formula for constructing words and non-words that implicitly conjure positive or negative emotion. Accordingly, we see potential applications of this study to a variety of social, educational, clinical, and marketing contexts. Consider a few examples. Advertising campaigns, whether for commercial products or political candidates, attempt to enhance emotional salience or appeal through carefully chosen language (Lerman & Garbarino, 2002; Robertson, 1989). Likewise, educational and mental health professionals seek to destigmatise certain conditions or activities through more sensitive labelling (Ben-Zeev, Young, & Corrigan, 2010; Siperstein, Pociask, & Collins, 2010). Even in artistic contexts, such as film and literature, these acoustic principles could be applied to evoke a particular emotional subtext. Indeed, our data suggest that "Darth Vader" (Lucas, 1977) is an acoustically more appropriate name for an intergalactic miscreant than "Barth Faber", by virtue of the downward frequency shifts and thus inherently negative emotional connotation. The results of this study elucidate an acoustic

mechanism by which the characteristic properties of phonemes may communicate emotion.

Manuscript received 20 March 2012
Revised manuscript received 27 November 2012
Manuscript accepted 28 November 2012
First published online 3 January 2013

REFERENCES

Bachorowski, J. A., & Owren, M. J. (1995). Vocal expression of emotion—Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science*, *6*(4), 219–224.

Bachorowski, J. A., & Owren, M. J. (1996). Vocal expression of emotion is associated with vocal fold vibration and vocal tract resonance. *Psychophysiology*, *33*, S20–S20.

Ben-Zeev, D., Young, M. A., & Corrigan, P. W. (2010). DSM-V and the stigma of mental illness. *Journal of Mental Health*, *19*(4), 318–327. doi:10.3109/09638237.2010.492484

Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, *66*, 1001–1017.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–298.

Charlton, B. D., Ellis, W. A., Larkin, R., & Fitch, W. T. (2012). Perception of size-related formant information in male koalas (*Phascolarctos cinereus*). *Animal Cognition*, *15*(5), 999–1006.

Charlton, B. D., Ellis, W. A., McKinnon, A. J., Cowin, G. J., Brumm, J., Nilsson, K., & Fitch, W. T. (2011). Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: Honesty in an exaggerated trait. *The Journal of Experimental Biology*, *214*(Pt. 20), 3414–3422.

Charlton, B. D., Reby, D., & McComb, K. (2008). Effect of combined source (F0) and filter (formant) variation on red deer hind responses to male roars. *Journal of the Acoustical Society of America*, *123*(5), 2936–2943.

Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., & Maneewongvatana, S. (2008). Encoding emotions in speech with the size code. A perceptual investigation. *Phonetica*, *65*(4), 210–230.

Davis, R. (1961). The fitness of names to drawings. A cross-cultural study in Tanganyika. *British Journal of Psychology*, *52*, 259–268.

Delattre, P., Liberman, A., & Cooper, F. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, *27*, 769–773.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*(2, Pt. 1), 1213–1222.

Fitch, W. T. (1999). Acoustic exaggeration of size in birds by tracheal elongation: Comparative and theoretical analyses. *Journal of Zoology*, *248*, 31–49.

Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*(7), 258–267.

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London, Series B*, *268*, 1669–1675.

Folkins, C., & Lenrow, P. (1966). An investigation of the expressive values of graphemes. *The Psychological Record*, *16*, 193–200.

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

Klink, R. R. (2000). Creating brand names with meaning: The use of sound symbolism. *Marketing Letters*, *11*(1), 5–20.

Kohler, W. (1929). *Gestalt psychology*. New York, NY: Liveright.

Lerman, D., & Garbarino, E. (2002). Recall and recognition of brand names: A comparison of word and nonword name types. *Psychology and Marketing*, *19*, 621–639.

Liberman, A., Delattre, P., Cooper, F., & Gerstman, L. (1954). The role of consonant–vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, *68*, 1–13.

Lucas, G. [Writer]. (1977). *Star Wars* [Film]. 20th Century Fox.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults [Research Support, Non-US Gov't]. *Developmental Science*, *9*(3), 316–322. doi:10.1111/j.1467-7687.2006.00495.x

Miron, M. (1961). A cross-linguistic investigation of phonetic symbolism. *Journal of Abnormal and Social Psychology*, *62*, 623–630.

Newman, S. (1933). Further experiments in phonetic symbolism. *American Journal of Psychology*, *45*, 53–75.

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought, and language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1566), 941–947. doi:10.1098/rspb.2004.2954 QWEU7P6HQ0-BE5QP1 [pii]

Repp, B. H., & Svastikula, K. (1988). Perception of the [m]–[n] distinction in VC syllables. *Journal of the Acoustical Society of America*, *83*(1), 237–247.

Riede, T., & Fitch, T. (1999). Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*). *The Journal of Experimental Biology*, *202*(Pt. 20), 2859–2867.

Robertson, K. (1989). Strategically desirable brand name characteristics. *Journal of Consumer Marketing*, *6*, 61–71.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*, 225–239.

Shrum, L. J., & Lowrey, T. M. (2007). Sounds convey meaning: The implications of phonetic symbolism for brand name construction. In T. M. Lowrey (Ed.), *Psycholinguistic phenomena in marketing communications* (pp. 39–58). Mahwah, NJ: Erlbaum.

Siperstein, G. N., Pociask, S. E., & Collins, M. A. (2010). Sticks, stones, and stigma: A study of students' use of the derogatory term "retard". *Intellectual and Developmental Disabilities*, *48*(2), 126–134.

Smith, D. R., Patterson, R. D., Turner, R., Kawahara, H., & Irino, T. (2005). The processing and perception of size information in speech sounds. *Journal of the Acoustical Society of America*, *117*(1), 305–318.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, *64*(5), 1358–1368.

Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M., & Kurowski, K. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America*, *91*(5), 2979–3000.

Sussman, H. M., Bessell, N., Dalston, E., & Majors, T. (1997). An investigation of stop place of articulation as a function of syllable position: A locus equation perspective. *Journal of the Acoustical Society of America*, *101*(5, Pt. 1), 2826–2838.

Taylor, A. M., Reby, D., & McComb, K. (2010). Size communication in domestic dog, *Canis familiaris*, growls. *Animal Behaviour*, *79*(1), 205–210.

Taylor, I. K., & Taylor, M. M. (1965). Another look at phonetic symbolism. *Psychological bulletin*, *64*(6), 413–427.

Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Quarterly Journal of Experimental Psychology*, *64*(12), 2392–2404.

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

Downloaded by [University of Wisconsin - Madison] at 08:53 03 January 2013