

LETTER TO THE EDITOR

A resource of validated digital audio recordings to assess identification of emotion in spoken language after a brain injury

BOAZ M. BEN-DAVID^{1,2,3,4}, ARTHIGA THAYAPARARAJAH^{2,4},
& PASCAL H. H. M. VAN LIESHOUT^{2,3,4,5,6}

¹Communication Aging and Neuropsychology Lab, School of Psychology, Interdisciplinary Center Herzliya (IDC), Herzliya, Israel, ²Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, Canada, ³Toronto Rehabilitation Institute, Toronto, Ontario, Canada, ⁴Department of Psychology, Centre for Research on Biological Communication Systems, University of Toronto Mississauga, Mississauga, Toronto, Ontario, Canada, ⁵Graduate Department of Rehabilitation Science, University of Toronto, Toronto, Ontario, Canada, and ⁶Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada

Background

A recent paper [1] presented a set of validated and linguistically equated lexical sentences that can test for the ability of persons with brain injury to process emotions in the lexical content. Here is presented a set of validated digital audio recordings of these sentences (see Electronic Appendix), spoken in various prosodies. This set can be used to separate the impact of lexical content and prosody on the processing of emotion in speech in persons with brain injury. This set is made available to researchers and clinicians.

A reduced ability to recognize emotional cues following an acquired brain injury (ABI) has a detrimental impact on rehabilitation. Specifically, emotion-identification difficulties (EID) in speech following ABI are found to be widespread [2] and related to poor quality-of-life [3]. As a consequence, the past decade shows a growing interest in both the research and clinical community to correctly assess EID as a path for rehabilitation. Yet, existing tools are scarce and may not fit an ABI population [1], as they cannot provide a complete picture of the specific roles of prosody (tone of speech) and lexical content in EID. For example, when one hears a sentence that carries a happy lexical content—‘I feel wonderful today’—spoken in an angry prosody,

what emotion was conveyed by the speaker: Anger, happiness or a combination of both? EID in ABI may result from an impaired identification of prosodic auditory cues, of the lexical content of a spoken utterance, or from an impaired ability to selectively attend to one dimension when the other also conveys emotional information.

In our previous work [1], we presented a set of 50 lexical sentences, with 10 exemplars in each of the following five categories: Anger, Fear, Happiness, Sadness and Neutral. These lexical sentences were distinctive in conveying their corresponding emotions and they were equated on main linguistic characteristics (e.g. frequency of usage, phonologic neighbourhood, sentence length) across the five emotion categories. This way, this lexical set controls for well-known biases on word and emotion identification. The next step in EID assessment is presented here, with a set of digital audio recordings of these lexical sentences spoken in five prosodies corresponding to the emotional categories mentioned above.

Stimuli selection

The 50 lexical sentences were recorded by a native English-speaking (Ontario, Canada) actress.

Table I. The 50 spoken sentences, as distributed across emotions expressed in prosody and lexical content.

Lexical content	Prosody				
	Anger	Fear	Sad	Happy	Neutral
Anger	1. You over-charged me for that. 2. Stop what you're doing and listen to me.	1. Stop what you're doing and listen to me. 2. Go to hell.	1. Quiet, this is a library. 2. Get out of my room.	1. Don't waste my time. 2. I am very angry.	1. Do not push your luck. 2. I'm sick of you being late.
Fear	1. I can hear foot-steps in the night. 2. I can hear a sharp scream from behind.	1. Watch out for that tiger. 2. Run for your life.	1. I can't see the bear but I can hear it. 2. I smell the gas leaking from the stove.	1. The fire is spreading to the gas pipe. 2. Look out there's a car coming.	1. It's about to explode. 2. He has a knife.
Sad	1. I'm going to a funeral. 2. This song makes me cry.	1. Gray clouds make me feel gloomy. 2. This scene makes him feel blue.	1. This is a sad moment. 2. I am so lonely.	1. I've been crying all day. 2. My best friend is moving away.	1. The weather is depressing. 2. My pet died today.
Happy	1. Congratulations, you're hired. 2. This is my favourite part.	1. Good job, the crowd really loves you. 2. I got promoted in my job.	1. Great, you got first place. 2. This is the happiest day of my life.	1. I won an award. 2. This food tastes very good.	1. It's a beautiful day outside. 2. I feel wonderful today.
Neutral	1. He stands on the deck. 2. Digital clocks are common.	1. Lots of bins are in the room. 2. A bag is in the room.	1. Four drawers are in the cabinet. 2. This is a garbage can.	1. Red pipes are metallic. 2. Our body is made of water.	1. I see a rug on the floor. 2. Containers have a blue lid.

Each sentence was recorded three times in each of the five different prosodies (Anger, Fear, Happiness, Sadness and Neutral) to generate a set of 750 recorded sentences. Digital audio files were equated with respect to root mean square amplitude. From these spoken sentences, a sub-set of 50 was selected such that no lexical sentence is repeated and each lexical emotional category is represented twice in each of the tested emotional prosodies, generating a 5×5 matrix (Table I). These spoken sentences were chosen based on high ratings from a group of six young adults, reflecting good exemplars of their respective emotional prosodies. This set of 50 spoken sentences is available in an Electronic Appendix.

Stimuli testing

To validate the prosody without the influence of the lexical content, two methods of digital acoustic filtering were employed: (a) prosody unveiling restricted representation (PURR [4]); and (b) low-pass filtering (LPF500 [5]). The former involves eliminating all frequencies above the third harmonic and aperiodic signals, while the latter simply removes all frequencies above 500Hz. Both are reported in the literature as effective techniques to de-lexicalize recorded speech, while leaving enough

acoustic information for the identification of prosodic cues.

Participants

Twenty-eight students from the University of Toronto (18–28 years old) participated in this study. All were native English speakers as assessed by a self-report and a vocabulary test (Mill-Hill), had pure-tone air-conduction thresholds within clinically normal limits in the 0.25–3 kHz range in both ears (≤ 20 dB HL) and no indication of head trauma or CNS diseases. Randomly, 12 participants were assigned to the PURR and 16 to the LPF500 condition.

Procedure

In both conditions, participants were tested individually in a single-walled sound-attenuating booth. The 50 acoustically manipulated files (PURR or LPF500) were presented binaurally via headphones, once in each of four rating blocks (in a pseudo-random fashion), corresponding to Anger-, Fear-, Sadness- and Happiness-rating blocks, making for 200 trials. In each rating block, participants were asked to rate how much they agree that the speaker identifies with the pre-defined emotion (based on the prosody) using a 6-point Likert scale ranging from strongly disagree (1) to strongly agree (6).

Results

Across emotional-rating blocks, sentences spoken in the corresponding rated-prosody (means, LPF500: 4.4; PURR: 4.1) received scores higher than the average values of other emotional prosodies, while neutral-prosody sentences (means, LPF500: 1.7; PURR: 2.6) received lower ratings than the average (linear trend, LPF500: $F(1, 16) = 365.5$, $p < 0.001$, $\eta_p^2 = 0.96$; PURR: $F(1, 11) = 80.5$, $p < 0.001$, $\eta_p^2 = 0.88$). For example, in the anger-rating block ratings given to acoustically manipulated anger-prosody sentences (means, LPF500: 4.5; PURR: 4.4) were higher than average values for Fear, Happy and Sad prosodies and ratings given to acoustically manipulated neutral-prosody sentences (means, LPF500: 1.9; PURR: 2.2) were lower than the average. In sum, these results confirm that listeners identified the intended emotional valence of emotional-prosody sentences and did not perceive any emotional valence in neutral-prosody sentences.

Discussion

In this letter we argue that, in order to obtain a more complete picture on EID, it is necessary to carefully investigate the specific influences of prosody and lexical content on the comprehension of spoken language. To this end, the Electronic Appendix includes a set of 50 validated and controlled, emotional and neutral, spoken sentences, available for use in research on emotion processing. Furthermore, since EID has significant impacts on the rehabilitation of ABI, we believe having access to this set can be useful to improve the assessment and treatment of communication problems related to emotion processing in this population. Given its general nature, this set can also be employed with

other populations (e.g. ageing, dementia) characterized by impaired communication skills with an underlying emotional processing deficit component.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Ben-David BM, Van Lieshout PHHM, Leszcz T. A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Injury* 2011;25:206–220.
2. Dimoska A, McDonald S, Pell M, Tate R, James C. Recognizing vocal expressions of emotion in patients with social skills deficits following traumatic brain injury. *Journal of the International Neuropsychological Society* 2010;16: 369–382.
3. Henry JD, Phillips LH, Crawford JR, Theodorou G, Summers F. Cognitive and psychosocial correlates of alexithymia following traumatic brain injury. *Neuropsychologia* 2006;44:62–72.
4. Sonntag GP, Portele T. PURR – A method for prosody evaluation and investigation. *Computer Speech and Language* 1998;12:437–451.
5. Knoll AM, Uther M, Costall A. Effects of low-pass filtering on the judgment of vocal affect in speech directed to infants, adults, and foreigners. *Speech Communication* 2009;51: 210–216.

Electronic Appendix

A folder containing 50 digital audio files (WAV) is available online. File names are compatible with Table I.