Brain and Cognition 96 (2015) 28–37

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

The effects of sad prosody on hemispheric specialization for words processing

Rotem Leshem [a,b,*], Yossi Arzouan [c], Rinat Armony-Sivan [d]

[a] Department of Psychology, University of California, Los-Angeles 90095-1563, USA
[b] Department of Criminology, Bar-Ilan University, Ramat Gan 52100, Israel
[c] Department of Education, Levinsky College of Education, Tel-Aviv 61481, Israel
[d] Department of Psychology, Ashkelon Academic College, Ashkelon 78211, Israel

ARTICLE INFO

Article history:
Accepted 16 March 2015
Available online 1 April 2015

Keywords:
Hemispheric specialization
Speech perception
Prosody
Emotion
Attention

ABSTRACT

This study examined the effect of sad prosody on hemispheric specialization for word processing using behavioral and electrophysiological measures. A dichotic listening task combining focused attention and signal-detection methods was conducted to evaluate the detection of a word spoken in neutral or sad prosody. An overall right ear advantage together with leftward lateralization in early (150–170 ms) and late (240–260 ms) processing stages was found for word detection, regardless of prosody. Furthermore, the early stage was most pronounced for words spoken in neutral prosody, showing greater negative activation over the left than the right hemisphere. In contrast, the later stage was most pronounced for words spoken with sad prosody, showing greater positive activation over the left than the right hemisphere. The findings suggest that sad prosody alone was not sufficient to modulate hemispheric asymmetry in word-level processing. We posit that lateralized effects of sad prosody on word processing are largely dependent on the psychoacoustic features of the stimuli as well as on task demands.

1. Introduction

Speech processing is considered a bilateral neurocognitive process in which each hemisphere is specialized for a different function (Poeppel, 2003). The left hemisphere (LH) is believed to deal predominantly with the linguistic aspects of speech (i.e., phonology and syntax), while the right hemisphere (RH) is focused on paralinguistic factors related to the speaker's voice and emotional state, including speech prosody and other emotional components (Buchanan et al., 2000; Hugdahl, 2000; Hugdahl, 2005; Poeppel, 2003; Springer & Deutsch, 1997). Thus, the RH specializes in sensitivity to patterns of pitch, loudness, and stimulus length, which signal the emotional state and communicate the emotional intentions of the speaker (Grimshaw, Séguin, & Godfrey, 2009; Witteman, van Ijzendoorn, van de Velde, van Heuven, & Schiller, 2011).

Despite this commonly accepted view, the scientific literature reveals a debate about the roles of the LH and RH in processing emotional prosody, particularly in the domain of emotional word processing (Eviatar & Zaidel, 1991; Smith & Bulman-Fleming, 2005), which is part of an ongoing debate about hemispheric specialization in perception of visual and auditory emotional stimuli (for review, see: Demaree, Everhart, Youngstrom, & Harrison, 2011). According to the "Right Hemisphere Model" the RH is specialized for the perception, expression, and experience of emotions, regardless of valence (Tucker, 1981). In contrast, the "Valence Model" posits that the RH is specialized for negative emotion and that the LH is specialized for positive emotion (Demaree et al., 2011). The Valence Model was largely subsumed by the approach-withdrawal model of emotion processing, which hypothesized that emotions associated with approach behaviors and withdrawal behaviors are processed within the left- and right-anterior brain regions, respectively (Davidson, 1995; Demaree et al., 2011).

In considering neural asymmetry in the processing of words presented in emotional prosody, it is necessary to go beyond these theories of emotional processing and examine the effects of hemispheric specialization based on psychophysical properties of the acoustic signal as well. Research has suggested that LH lateralization for speech is a secondary consequence of LH specialization for processing rapidly changing acoustic information such as that involved in speech perception (Schwartz & Tallal, 1980). Collectively known as cue-dependent hypotheses (Gandour et al., 2004; Zatorre & Belin, 2001), these theories propose that there are differences in hemispheric recruitment deriving from the

* Corresponding author at: Department of Criminology, Bar-Ilan University, Ramat Gan 52100, Israel.
E-mail address: rotem.leshem@biu.ac.il (R. Leshem).

http://dx.doi.org/10.1016/j.bandc.2015.03.002
0278-2626/

manner in which information relevant to speech is quantized in the time domain. For example, according to the Asymmetric Sampling in Time (AST) hypothesis put forward by Poeppel (2003), temporal features (shorter integration times) are processed predominantly in the LH and spectral features (longer temporal windows) in the RH (Boemio, Fromm, Braun, & Poeppel, 2005; Gandour et al., 2002; Hickok & Poeppel, 2007; Nicholls & Lindell, 2000; Schirmer & Kotz, 2006; Zatorre, Bouffard, Ahad & Belin, 2002). This model predicts that syllable-initial stop consonants (e.g., ba/da) cued by rapid up- or down-going shifts in spectral energy distribution (formant transitions) extending across a few tens of milliseconds (Hertrich, Mathiak, Lutzenberger, & Ackermann, 2002) will preferentially recruit the LH. Indeed, fast temporal/spectral changes produce a reliable LH advantage when measured in dichotic listening to consonant–vowel syllable or consonant–vowel-consonant syllable paradigms, indicating higher proficiency of the LH in processing formant transitions (Bryden, 1988; Hertrich et al., 2002; Schwartz & Tallal, 1980; Tervaniemi & Hugdahl, 2003). Functional asymmetry in temporal processing has been consistently confirmed by the discovery of LH specialization for the analysis of rapidly changing acoustic cues in consonant–vowel syllables (Eichele, Nordby, Rimol, & Hugdahl, 2005; Hertrich et al., 2002; Hugdahl, 2005; Jäncke & Shah, 2002; Sandmann et al., 2007; Schwartz & Tallal, 1980; Tallal, Miller, & Fitch, 1993), and of increased activation in the RH for slow, compared to rapid, acoustic transitions in speech and non-speech sounds (Boemio et al., 2005; Schirmer & Kotz, 2006; Zatorre & Belin, 2001).

The AST hypothesis holds that left auditory areas preferentially extract information from short temporal integration windows (∼20–50 ms) while homologous areas in the right hemispheric extract information from long integration windows (∼150–250 ms). Thus, this model predicts that segmental information extractable in a short time window will be processed primarily by the LH whereas prosodic information extractable only in long time windows will be processed by the RH (Poeppel, 2001; Poeppel, 2003). Moreover, ERP studies have confirmed an early time course for syllable- and word-level processing in the time range of 100–150 ms, associated with the N1 component, whereas emotional word effects have been demonstrated in the time range of 180–300 ms post-stimulus, associated with the P2 component (Scott, O'Donnell, Leuthold, & Sereno, 2009). This supports the possibility that RH plays a greater role in processing words presented in emotional prosody than in processing neutral words, especially in later stages of processing (Hickok & Poeppel, 2007).

It should also be noted that the acoustic markers of vocal prosody such as fundamental frequency and speech rate each have contributions to the effects of emotional prosody on linguistic processing. Fundamental frequency or pitch frequency (F0), which corresponds to vocal cord vibrations for vocalic sounds of speech, is the main acoustic cue for intonation and stress in speech, and is crucial in tone languages for phoneme identification (Deng & O'Shaughnessy, 2003). Emotionally depressed, sad, or ashamed speakers produce speech with very little variation in F0 and a slow speech rate, while excited emotional state in the speaker, such as surprise, interest, joy, contempt, or anger results in increased variation in F0 and fast speech rate (Scherer, 2003). These differences may also affect hemispheric asymmetry in processing emotional speech.

Given the wide range of factors believed to affect hemispheric asymmetry in emotional word processing, the purpose of the present study was to further examine the effects of emotional prosody on hemispheric asymmetry for word processing using both behavioral and electrophysiological measures. Specifically, we examine whether emotional prosody can modulate hemispheric asymmetry for word processing when the word is task-relevant and prosody is task-irrelevant.

To evaluate hemispheric specialization in the context of prosodic processing of words and emotions, Bryden and MacRae (1988) introduced a dichotic listening test known as "Dichotic-Listening to Words and Affects" (DLWA). The task consists of four dichotically-paired words (bower, dower, power, tower) spoken in four different emotional tones (sad, happy, angry, neutral). Participants are required to focus their attention on one ear at a time and identify either a target word (word task) or the target affect/emotion (emotion task). Through this procedure, the DLWA task allows researchers to examine whether a word or emotion was processed in a single hemisphere or bilaterally. Bryden and MacRae (1988) reported a significant right ear advantage (REA) in the word task and left ear advantage (LEA) in the emotion task. The REA in word processing was consistent with previous studies on language processing. However, the relationship between the LEA and emotional word processing had not previously been reported. It was interpreted as providing support for the RH hypothesis, although a trend in the results of this study toward greater LEAs for emotional tones with negative valence could be construed as partial evidence for the valence hypothesis.

Grimshaw, Kwassny, Covell, and Johnson (2003) used the same DLWA task described above, but focused mainly on neutral and sad emotional tones. The main finding of their study was that in the word task, the REA typically observed when a word was spoken in neutral prosody was attenuated when it was spoken in a sad emotional prosody. The researchers proposed that this effect of sad prosody on word processing reflected RH involvement in emotional word processing. Several other behavioral studies using a similar paradigm showed comparable results (Grimshaw et al., 2009; Hale, Zaidel, McGough, Phillips, & McCracken, 2006). Grimshaw et al. (2009) replicated and extended this work, comparing the effects of neutral prosody to those of happy, angry, and sad prosodies on word processing. Results showed that unlike the decreased REA effect on words spoken in sad prosody, angry and happy prosody had no effect. They argued that emotional prosody itself did not facilitate RH linguistic processing and that there are might be other factors such as psychoacoustic properties that need to be considered.

In the present study we employed the DLWA word task version, as introduced by Grimshaw et al. (2009) designed specifically to examine the role of LH and RH in detecting a word spoken in neutral or emotional prosody, when attending to either the right or the left ear. Using ERP measures enable examination of the involvement of LH and RH in the processing of emotionally prosodic words when attending to either the right or the left ear. To this end, based on the review above and following the studies of Grimshaw et al. (2003), Grimshaw et al. (2009), 'sad' prosody was utilized. To gain some degree of experimental control over attentional bias for LH in word processing (Jerger & Martin, 2004), we combined instructions to focus attention on one ear at a time for a whole block of trials (blocked design) and to detect a specific verbal stimulus spoken in neutral or sad prosody in each block. As such, the task involved two levels of cognitive information processing, stimulus-driven/bottom-up processing, which is related to detection of the word, and instruction-driven/top-down processing, which is related to directing attention toward a specific ear (Hiscock, Inch, & Ewing, 2005; Hugdahl, 2005; Leshem, 2013; Thomsen, Rimol, Ersland, & Hugdahl, 2004). Using both focused attention and signal-detection procedures, the current paradigm has the advantage of probing each hemisphere separately. As reviewed above, word detection is associated with a REA/LH dominance. Therefore, word detection when attending to the left ear may be more complex, due to interference between stimulus-driven/bottom-up and instruction-