Journal of Neurolinguistics 22 (2009) 313–326

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

# Once more with feeling: The effects of emotional prosody on hemispheric specialisation for linguistic processing

Gina M. Grimshaw*, Julie Anne Séguin, Hazel K. Godfrey

*Victoria University of Wellington, Wellington, New Zealand*

ARTICLE INFO

*Article history:*
Received 13 May 2008
Received in revised form 23 October 2008
Accepted 27 October 2008

*Keywords:*
Hemispheric specialisation
Prosody
Emotion
Language
Speech perception

ABSTRACT

Speech perception requires a bilateral network of neural systems, with the left hemisphere specialised for linguistic processes including phonology, and the right hemisphere specialised for paralinguistic processes including emotional prosody. The present study used dichotic listening to determine how these two processing systems interact. Experiment 1 confirmed that the left hemisphere was specialised for the linguistic task, and the right hemisphere was specialised for the prosodic task. In Experiment 2, participants performed the linguistic task, but the words were spoken in prosodies that were neutral, happy, angry, or sad. Although the emotional prosody was irrelevant for the task, there was an attenuation of the typical right ear advantage when words were spoken with sad prosody, suggesting greater right hemisphere contributions to linguistic processing for sad speech. Similar effects were not observed with angry or happy prosody, suggesting that emotional prosody *per se* does not facilitate right hemisphere linguistic processing. Results are interpreted in terms of both psychoacoustic and emotion-specific theories of hemispheric specialisation.



## 1. Introduction

Although the special role of the left hemisphere in speech processing has been acknowledged for over a century, the right hemisphere also plays an important role. In fact, speech processing is perhaps best considered a bilateral system. The left hemisphere is specialised for linguistic aspects of speech

* Corresponding author. School of Psychology, Victoria University of Wellington, PO Box 600, Wellington, New Zealand. Tel.: +64 04 463 6420; fax: +64 04 463 5420.

*E-mail address:* gina.grimshaw@vuw.ac.nz (G.M. Grimshaw).

0911-6044/$ – see front matter 
doi:10.1016/j.jneuroling.2008.10.005

processing such as phonology and syntax; whereas the right hemisphere is specialised for paralinguistic aspects, including information that is carried in the voice such as the speaker's emotional state, attitudes, sex, and age (Friederici & Alter, 2004; Lattner, Meyer, & Friederici, 2005; Poeppel, 2003). One important right hemisphere contribution to speech perception is the comprehension of emotional prosody; the patterns of pitch, loudness, and length that signal the emotional state and communicate the emotional intentions of the speaker. Studies of brain damaged patients (Borod et al., 1998; Bowers, Coslett, Bauer, Speedie, & Heilman, 1987; Charbonneau, Scherzer, Aspirot, & Cohen, 2003; Kucharska-Pietura, Phillips, Gernand, & David, 2003; Walker, Daigle, & Buzzard, 2002), perceptual asymmetries in normals (Bryden & MacRae, 1989; Grimshaw, 1998; Grimshaw, Kwasny, Covell, & Johnson, 2003; Stirling, Cavill, & Wilkinson, 2000), and neuroimaging research (Buchanan et al., 2000; Gandour et al., 2003; Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003) all indicate a primary role for the right hemisphere in the comprehension of emotional prosody (for a comprehensive review see Borod, Zgaljardic, Tabert, & Koff, 2001). The specific neural substrates of prosodic processing are still a matter of some debate (Berckmoes & Vingerhoets, 2004; Pell, 2006; Van Lancker & Sidtis, 1992), but cortical contributions appear to involve a distributed network of right frontal and temporal structures (for a review see Schirmer & Kotz, 2006).

While coherent speech comprehension clearly involves the integration of linguistic and prosodic levels of analysis, very little is known about the neural basis of linguistic-prosodic interaction. Although the left hemisphere is clearly specialised for linguistic processing, such lateralisation is relative and not absolute. Converging evidence from studies of both patients and neurologically intact individuals indicates that the right hemisphere has considerable linguistic as well as paralinguistic competence (Beeman & Chiarello, 1998; Gainotti, Caltagirone, & Masullo, 1981; Gazzaniga & Sperry, 1967; Gold & Kertesz, 2000; Zaidel, 1976). It is therefore possible that the right hemisphere may play a greater role in processing prosodically-emotional speech than in processing neutral speech.

One approach to this question is to determine whether the presence of emotional prosody modulates hemispheric contributions to linguistic processing. Grimshaw et al. (2003) examined this question using a dichotic listening task in which two words were presented simultaneously, one to each ear. Participants listened for a target word (a linguistic task) when both words were spoken in either a neutral or a sad prosody. Participants were directed to pay attention to the word and ignore the tone of voice in which it was spoken. Linguistic tasks such as this typically produce a right ear advantage (REA), reflecting left hemisphere specialisation for linguistic processing (Bryden, 1988; Kimura, 1961). As expected, a robust REA was observed when words were spoken in a neutral tone of voice. However, when words were spoken in a sad tone of voice, the REA was attenuated in the accuracy data, and eliminated entirely in the response time data. They hypothesized that the presence of the sad emotional prosody activated the right hemisphere sufficiently to facilitate right hemisphere linguistic processing (Grimshaw et al., 2003).

However, Grimshaw et al. (2003) was limited to a comparison of effects of neutral and sad prosody. The present study replicated and extended this work to compare the effects of happy, angry, and sad emotional prosody on hemispheric specialisation for a linguistic task. This comparison allows us to determine whether the effect of sad prosody on linguistic processing extends to other emotional prosodies, or is specific to sadness. The right hemisphere specialisation for emotional prosodic comprehension is observed regardless of emotional valence (Borod et al., 1998; Bryden & MacRae, 1989; Buchanan et al., 2000). Thus, if the alteration in linguistic processing is driven by right hemisphere prosodic activation, then it should be observed for all emotional prosodies.

It is possible, however, that the prosodic modulation of linguistic processing is specific to sad prosody. Emotional prosodies differ at the acoustic level; sad prosodies have a low amplitude, a slow speech rate, and little variability in fundamental frequency ($F_0$), while happiness is characterized by a fast speech rate, and both angry and happy prosodies are higher in $F_0$ and $F_0$ variability (for a review, see Scherer, 2003). It is possible that the acoustic properties of sadness facilitate right hemisphere linguistic processing. For example, Poeppel (2003) has argued that the left and right hemispheres differ in temporal resolution, with the left hemisphere sampling over windows of 20–40 ms (well-suited to phoneme discrimination) whereas the right hemisphere samples over windows of 150–250 ms (optimal for extraction of prosodic contours). Possibly the slower speech associated with sad prosody allows for phoneme extraction at the right hemisphere's sampling rate.

Alternatively, differences might be observed between sad and other prosodies for emotional reasons. Although the right hemisphere is specialised for the comprehension of all emotional prosodies, there are hemispheric differences in emotional experience (for a review see Borod et al., 2001). It is very unlikely that listening to emotional prosodies in a target detection task could induce emotional experience; however, it is possible that emotional prosody might activate networks that are associated with emotional experience. According to Heller's circumplex model of emotion (Heller, 1990, 1993), greater left than right frontal activity is associated with positive emotion and greater right than left frontal activity is associated with negative emotion. If the effect of emotional prosody on linguistic processing depends on valence, then one would expect attenuation of the ear advantage associated with sad and angry prosody (negative valence), but not with happy prosody (positive valence). An alternative hypothesis is that the hemispheres differ in approach/withdrawal motivation (Davidson, 1992, 1993; Davidson, Ekman, Saron, Senulis, & Friesen, 1990), with greater left than right frontal activity related to approach motivations (e.g., those associated with happiness and anger), and greater right than left frontal activity related to withdrawal motivations (e.g., those associated with sadness and fear). If the emotional prosody effect is related to motivational drive then one would expect it to be observed for sad prosody, but not happy or angry prosody.

Although there are no reports in the literature on effects of emotional prosody on hemispheric asymmetry for linguistic processing (beyond that in Grimshaw et al. 2003), there is some evidence that language lateralisation can be modulated through right hemisphere emotional activation. For example, Gadea, Gómez, González-Bono, Espert, and Salvador (2005) found an attenuation of the REA for a consonant-vowel (CV) discrimination task with a negative mood induction. Using a similar dichotic task, Asbjörnsen, Hugdahl, and Bryden (1992) found an attenuation of the REA in participants who were threatened with electric shock (presumably a fear-inducing condition), but not in participants who were promised high financial rewards (presumably inducing a positive affect). In a study that is perhaps most analogous to the prosody manipulation of interest here, Van Strien and Boon (1997) had participants do a lateralised lexical decision task while simultaneously listening to noise, positive music or aversive music. They found that the right visual field advantage (RVFA) that is typically observed for lexical decision was eliminated in the aversive music condition, but not in the noise or positive music condition. Furthermore, the aversive music had an effect only on words, not non-words, suggesting that the effect reflected a change in hemispheric asymmetry for language processing, and not just a leftward attentional bias induced by right hemisphere activation. Taken together, these findings suggest that language lateralisation may be modulated through emotional manipulation.

We report here two experiments to examine the effects of emotional prosody on lateralisation of a linguistic task. In Experiment 1, a dichotic listening procedure was used to examine hemispheric asymmetries in processing the linguistic and prosodic dimensions of speech. The primary aim of this experiment was to ensure that the stimuli produced the desired qualities of a REA (reflecting left-hemisphere specialisation) for the linguistic task (listening for a target word) and a LEA (reflecting right-hemisphere specialisation) for the prosodic task (listening for a target tone of voice). Furthermore, Experiment 1 allowed us to determine whether there were differences in the degree to which different emotional prosodies were lateralised to the right hemisphere. In Experiment 2, participants performed only a linguistic task, while listening to words that were spoken either with emotional prosody or in a neutral tone of voice. Following Grimshaw et al. (2003) it was expected that the ear advantage on the linguistic task would be attenuated in the sad prosody condition relative to the neutral prosody condition, reflecting an increase in right hemisphere contributions to linguistic processing. Experiment 2 allowed us to determine whether such effects are specific to sad prosody, are observed for all prosodies, or are influenced by prosodic valence.

## 2. Experiment 1

Experiment 1 was a replication of the classic study of Bryden and MacRae (1989) to ensure that our newly-constructed stimuli produced an REA when participants listened for a target word (the linguistic task) and a LEA when they listened for a target tone of voice (the prosodic task). The Bryden and MacRae stimuli were recorded by a North American male speaker; in using a New Zealand sample we wished to produce a set of stimuli spoken with the native regional accent.

## 2.1. Method

### 2.1.1. Participants

Thirty-two right-handed undergraduate psychology students (8 men and 24 women; mean age 18.91 years) who reported having no hearing deficits were recruited from a New Zealand university. All were native English speakers and received course credit for their participation. Handedness was assessed using the Waterloo Handedness Questionnaire – Revised (Elias, Bryden, & Bulman-Fleming, 1998); this questionnaire requires participants to indicate which hand they use for 36 skilled and unskilled activities. All participants wrote with their right hand, had no history of hand preference change, and had a positive (right-handed) score on the handedness questionnaire.

### 2.1.2. Stimuli and apparatus

The stimuli were the words "*bower*", "*dower*", "*power*", and "*tower*" spoken with a New Zealand accent by an adult male in prosodies that were neutral, angry, happy, and sad. These were the same words used by Bryden and MacRae (1989) and were chosen to be rhyming words (for better dichotic fusion) with two syllables (to better allow for the production of emotional prosody). The words were recorded using a Rode NT4 microphone and an M-Audio FireWire 410 mobile recording interface, controlled by an Apple MacBook. Sound files were first recorded using only one channel (mono) at 16 bits and 44.1 kHz and subsequently duplicated, using the recording and editing software Audacity, to create stereo files in which the left and right channels were identical. The digital stimuli were matched for peak amplitude and had a mean duration of 557 ms (range 469–647 ms).

Each auditory stimulus was analyzed with PRAAT (Boersma & Weenink, 2007) to describe its acoustic properties. The $F_0$ of the waveform of each of the 16 stimuli was extracted using the auto-correlation method in which speech waves at one time period are correlated with those in the consecutive time period (PRAAT uses a time period of 40 ms for analysis). The auto-correlation method uses this characteristic of speech wave forms to estimate the $F_0$ (see Johnson, 1997 for a more detailed description of the auto-correlation method). Only frequency values between 75 and 350 Hz (the appropriate range for a male speaker) were used for extraction of the $F_0$ (Van Lieshout, 2004). This condition was also imposed to eliminate the effect of the burst of air from the initial plosive, which for some of the stimuli was misinterpreted by PRAAT as a high frequency sound. Once the $F_0$ was extracted, the query function of PRAAT was used to calculate the mean, median, and standard deviation of each of the stimuli's $F_0$. Mean values of each parameter, collapsed across stimuli (bower, dower, power, and tower) are presented in Table 1.

A second acoustic analysis compared the speech rates of the stimuli. Because these stimuli differed only in the first phoneme, the analysis involved estimation of the duration of the initial plosive. The lengths of phonemes within words are necessarily subjective, as they are affected by coarticulation of surrounding sounds, and there are no clear physical boundaries between them. We had two raters estimate the duration of each plosive, using a combination of auditory playback and visual inspection of the sound wave. Each rater produced 5 independent estimates for each stimulus during different sessions. The shortest and longest estimates were then dropped, and a mean duration was calculated from the remaining three estimates. Inter-rater reliability based on these estimates was $r(16) = .838$. These values are presented in Table 1.

In a preliminary study, 10 participants (drawn from the same participant pool) listened to 64 recordings (four samples of each word/prosody combination) and were asked to identify the word as

**Table 1**
Acoustic parameters of stimulus items.

| Emotion | Mean $F_0$ (Hz) | Median $F_0$ (Hz) | SD $F_0$ (Hz) | Plosive duration (ms) |
|---|---|---|---|---|
| Happy | $195^a$ | $179^c$ | $75^a$ | $57^a$ |
| Angry | $184^a$ | $195^c$ | $45^a$ | $63^a$ |
| Sad | $123^b$ | $113^d$ | $25^b$ | $94^b$ |
| Neutral | $104^b$ | $103^d$ | $3^c$ | $50^c$ |

Note. Letters a,b,c within each column indicate means that do not differ from each other on post-hoc comparisons (Tukey's HSD).

well as the emotional tone and the intensity (on a scale of 1–5) with which it was spoken. The stimuli were randomly presented and repeated four times. The final stimulus set was constructed by selecting the 16 sound files (one for each word/tone combination) for which words were identified with a minimum accuracy of 92.5% ($M = 97.97\%$ SD = .03%), and emotions were identified with a minimum accuracy of 70% ($M = 87\%$, SD = .03%). Mean intensity ratings did not differ for happy ($M = 3.68$, SD = .55), angry ($M = 3.61$, SD = .75) or sad ($M = 3.41$, SD = .69), $F(2, 18) = .740$, $p = .491$, however, all three differed from neutral ($M = 1.88$, SD = .62), $F(1, 9) = 63.25$, $p < .001$.

Stimuli were presented using a Dell PC running Psychology Software Tools' E-Prime Suite version 1.0 (Schneider, Eschman, & Zuccolotto, 2002) through a pair of Manhattan noise-cancelling stereo headphones with circumaural cushions. We verified that stimuli were presented at a peak amplitude of 75 dB using a Philips sound meter. Headphone placement was counterbalanced across subjects to control for mechanical effects.

#### 2.1.3. Procedure

On each dichotic trial, participants heard two stimuli simultaneously; one in the right ear and one in the left ear. Participants heard all possible pairings of words and prosodies with the constraint that a different word and a different prosody were presented to each ear on each trial, producing a total of 144 unique stimulus pairs per block. Participants monitored for a specified target word (the linguistic task) for two blocks of trials, and for a specified target tone of voice (the prosodic task) for two blocks of trials, yielding a total of 576 trials. They indicated their response with a button press using the index (for present) or middle finger (for absent) of the specified hand. Response hand was alternated across blocks; half the participants listened for the word target first, and half listened for the prosody target first. Both conditions were preceded by 32 practice trials in which every word/tone combination was presented binaurally, twice and in random order, while participants monitored for their target. Participants were instructed to respond as quickly and as accurately as possible. Reaction time and accuracy were recorded using E-prime and the computer's internal timer. The experiment lasted approximately 1 h.

#### 2.1.4. Design

The basic design was a 2 (Task: linguistic, prosodic) × 2 (Ear: left, right) within-subjects design. Embedded within each task were 4 targets (*bower*, *dower*, *power*, and *tower*, for the linguistic task; *happy*, *angry*, *sad*, and *neutral* for the prosodic task) that were manipulated between-subjects.

### 2.2. Results and discussion

Hit rates, false alarm rates and mean response times for hits (correctly identifying the presence of a target) were calculated for each condition. Hit rates and false alarm rates (based on the assumption that false alarm rates were the same for the right and the left ears) were used to calculate $d'$, a measure of sensitivity, for each condition. An outlier procedure eliminated response times that were less than 200 ms (as anticipatory responses) and more than 2 standard deviations above a participant's mean. An initial omnibus Analysis of Variance (ANOVA) revealed no effects of response hand or task order on either response time or accuracy, and so they were eliminated from further analyses. Response time and accuracy measures are presented in Table 2.

#### 2.2.1. Accuracy analyses

Sensitivity, as measured by $d'$, was analyzed in a 2 (Task: linguistic, prosodic) × 2 (Ear: left, right) repeated measures ANOVA. A main effect of Task was observed, $F(1, 31) = 13.029$, MSE = .785, $p = .001$, with participants demonstrating better discrimination on the prosodic task ($M = 2.56$, SD = .83) than on the linguistic task ($M = 1.99$, SD = .61). A main effect of Ear, $F(1, 31) = 16.335$, MSE = .110, $p < .001$, was qualified by a Task × Ear interaction, $F(1, 31) = 28.277$, MSE = .119, $p < .001$. There was a significant REA for the linguistic task, $t(31) = 5.362$, $p < .001$, and a non-significant LEA for the prosodic task, $t(31) = 1.524$, $p = .138$.

In order to determine whether different targets produced different degrees of lateralisation, $d'$ values on the linguistic and prosodic tasks were analyzed in separate 2 (Ear) × 4 (Target) repeated

**Table 2**
Performance in Experiment 1 as a function of task and ear.

| | Left ear | | Right Ear | | Ear difference |
|---|---|---|---|---|---|
| | *M* | SD | *M* | SD | Mean |
| Linguistic task | | | | | |
| RT (ms) | 941 | 139 | 870 | 127 | 71 |
| Hit Rate | 0.63 | 0.16 | 0.79 | 0.14 | 0.16 |
| *d*′ | 1.72 | 0.56 | 2.28 | 0.64 | 0.56 |
| Prosodic task | | | | | |
| RT (ms) | 984 | 131 | 1013 | 156 | −29 |
| Hit Rate | 0.78 | 0.17 | 0.76 | 0.16 | −0.02 |
| *d*′ | 2.60 | 0.85 | 2.52 | 0.80 | −0.08 |

measures ANOVAs. No main effects or interactions involving target were observed, suggesting similar left hemisphere specialisation for the four linguistic targets, and similar right hemisphere specialisation for the four prosodic targets.

### 2.2.2. Response time analyses

Mean response times in each condition were analyzed in a 2 (Task) × 2 (Ear) repeated measures ANOVA. The main effect of Task, $F(1, 31) = 13.34$, MSE = 20,920, $p = .001$ reflected slower response times on the prosodic task ($M = 999$ ms, SD = 143 ms) than on the linguistic task ($M = 906$ ms, SD = 133 ms). This contrasts with the accuracy analysis, which showed higher accuracy on the prosodic task than on the linguistic task. However, this difference does not likely reflect a speed-accuracy trade-off, but rather the fact that the prosodic judgment, which is based largely on the pitch contour of the word, requires more information than the linguistic judgment, which requires attention to only the first phoneme. The main effect of Ear, $F(1, 31) = 4.899$, MSE = 2854, $p = .034$, was qualified by the Task × Ear interaction, $F(1, 31) = 22.436$, MSE = 3546, $p < .001$. As expected, a significant REA was observed for the linguistic task, $t(31) = 4.696$, $p < .001$, and a significant LEA was observed for the prosodic task, $t(31) = 2.202$, $p = .035$.

Effects of specific linguistic and prosodic targets were analyzed in separate Ear × Target ANOVAs for the linguistic and prosodic tasks. No main effects or interactions involving target approached significance, indicating that the left hemisphere was specialised for all linguistic targets equally, and that the right hemisphere was specialised for all prosodic targets equally.

This experiment confirmed that this new set of voice recordings, in a New Zealand accent, produced the expected pattern of hemispheric processing. Specifically, a REA (reflecting left hemisphere specialisation) was observed for the linguistic task, and a LEA (reflecting right hemisphere specialisation) was observed for the prosodic task. There were no differences among emotions on the prosodic task; all were equally lateralised to the right hemisphere. This pattern of results is consistent with other research showing a general right hemisphere specialisation for the perception of emotional prosody, regardless of valence (Borod et al., 1998; Bryden & MacRae, 1989; Buchanan et al., 2000; cf., Erhan, Borod, Tenke, & Bruder, 1998).

## 3. Experiment 2

The purpose of Experiment 2 was to examine the effect of specific emotional prosodies on lateralisation for linguistic processing. Participants performed a linguistic task for one block of trials in which the words were all spoken in a neutral tone of voice, and for one block of trials in which the words were all spoken in one emotional tone of voice (happy, angry, or sad). If activation associated with prosodic processing facilitates right hemisphere linguistic processing, then we expect that speech in all three prosodies will produce a reduced REA (or even an LEA) for the linguistic task compared to speech in a neutral prosody. Alternatively, right hemisphere linguistic facilitation may be specific to sad prosody, and will not be observed for happy or angry prosody.

### 3.1. Method

#### 3.1.1. Participants

Ninety-six right-handed undergraduate students (34 men and 62 women; mean age = 19.0 years) were randomly assigned to one of three emotional prosody conditions: *happy*, *angry*, or *sad*. Men and women were distributed evenly across groups. All were native English speakers and received course credit for their participation. Handedness was assessed using the Waterloo Handedness Questionnaire – Revised (Elias et al., 1998); all participants wrote with their right hand and had a positive (right-handed) score on the handedness questionnaire. The three emotion groups did not differ in mean age, $F(2, 93) = 0.136$, ns, or handedness index, $F(2, 93) = 1.285$, ns.

#### 3.1.2. Stimuli and apparatus

The computer apparatus and digital sound files were the same as those used in Experiment 1. Each trial consisted of the presentation of a word to the left ear while a different word was simultaneously presented to the right ear. However, in contrast to Experiment 1, both words were spoken in the same prosody. There were thus 12 unique pairs of stimuli spoken in each prosody, which were repeated three times to form blocks of 36 trials each.

#### 3.1.3. Procedure

Participants monitored for a target word (*bower*, *dower*, *power*, or *tower*) throughout the experiment (that is, they completed the linguistic task of Experiment 1). They completed two blocks (one with the left hand and one with the right hand) in which the words were spoken in a neutral tone of voice, and two blocks in which the words were spoken in an emotional tone of voice (happy, angry or sad, depending on group assignment). Block order was counterbalanced across participants. Each voice condition was preceded by 12 binaural practice trials. Participants were asked to respond as quickly and as accurately as possible. Response time and accuracy were recorded using E-prime and the computer's internal timer. Participation took approximately 40 min.

#### 3.1.4. Design

The experiment had a 2 (Prosody: neutral, emotional) × 2 (Ear: left, right) × 3 (Emotion Group: happy, angry, sad) mixed design, with Prosody and Ear as within-subject variables, and Emotion Group as a between-subjects variable.

### 3.2. Results and discussion

Hit rates, false alarm rates, and mean response times for hits were calculated for each condition (see Table 3). The signal detection measure $d'$ was calculated as a measure of sensitivity, based on the assumption that false alarm rates were the same for the right and the left ears. Response times of less than 200 ms or greater than two standard deviations above each participant's mean were eliminated as outliers. Preliminary omnibus ANOVAs revealed no effects of response hand or block order, and so they were eliminated from further analyses.

#### 3.2.1. Accuracy analyses

Sensitivity as measured by $d'$ values was analyzed in a 2 (Prosody: neutral, emotional) × 2 (Ear: left, right) × 3 (Emotion Group: happy, angry, sad) mixed model ANOVA, with Prosody and Ear as within-subject variables and Emotion as a between-subjects variable. The main effect of Prosody, $F(1, 93) = 9.952$, MSE = .922, $p = .002$, reflected better discrimination for words spoken in an emotional prosody ($M = 2.59$, SD = 1.04) than for words spoken in a neutral prosody ($M = 2.28$. SD = .99). The main effect of Ear, $F(1, 93) = 52.341$, MSE = .506, $p < .001$, was qualified by a three-way interaction of Prosody × Ear × Emotion Group, $F(2, 93) = 5.598$, MSE = .181, $p = .005$ (see Fig. 1). This interaction was further examined by running separate Prosody × Ear ANOVAs for each emotion. For participants who heard the sad prosody, a main effect of Ear, $F(1, 31) = 19.712$, MSE = .521, $p < .001$ was qualified by a Prosody × Ear interaction, $F(1, 31) = 11.892$, MSE = .197, $p = .002$. The REA for words spoken in the sad prosody ($M = .30$, SD = .74) was attenuated relative to the REA for words spoken in the neutral prosody

**Table 3**
Performance in Experiment 2 as a function of emotion group, prosody, and ear.

| Emotion group | Prosody | | | | | |
|---|---|---|---|---|---|---|
| | Neutral | | | Emotional | | |
| | LE | RE | Difference | LE | RE | Difference |
| RT (ms) | | | | | | |
| Happy ($N = 32$) | 837 (128) | 809 (129) | 28 (77) | 860 (180) | 829 (170) | 31 (118) |
| Angry ($N = 32$) | 868 (184) | 830 (168) | 38 (73) | 898 (184) | 871 (179) | 27 (91) |
| Sad ($N = 32$) | 844 (150) | 782 (108) | 62 (126) | 867 (124) | 849 (110) | 18 (68) |
| Mean | 850 (155) | 807 (137) | 43 (95) | 875 (164) | 849 (155) | 25 (93) |
| $d'$ | | | | | | |
| Happy ($N = 32$) | 2.09 (0.92) | 2.66 (0.75) | 0.57 (0.88) | 2.45 (1.15) | 3.08 (0.98) | 0.63 (0.80) |
| Angry ($N = 32$) | 2.02 (1.08) | 2.38 (1.09) | 0.36 (0.82) | 2.25 (0.96) | 2.70 (0.93) | 0.45 (0.77) |
| Sad ($N = 32$) | 1.86 (1.17) | 2.69 (0.91) | 0.83 (0.94) | 2.39 (1.08) | 2.69 (1.12) | 0.30 (0.74) |
| Mean | 1.99 (1.05) | 2.58 (0.93) | 0.59 (0.90) | 2.36 (1.06) | 2.82 (1.02) | 0.46 (0.77) |

Note. Standard deviations appear in (). LE, Left Ear; RE, Right Ear. For differences between ears, positive values reflect a right ear advantage, and negative values reflect a left ear advantage.

($M = .83$, SD = .94). In contrast, participants who heard the happy prosody demonstrated only a significant REA, $F(1, 31) = 22.133$, MSE = .524, $p < .001$, that did not interact with Prosody, F(1, 31) = .132, ns. Similarly, participants who heard the angry tone of voice also demonstrated a significant REA, $F(1, 31) = 11.227$, $p = .002$, that also did not interact with Prosody, $F(1, 31) = .435$, ns. In addition, those in the angry condition had better discrimination of their target when the words were spoken in an angry prosody than in a neutral prosody, $F(1, 31) = 6.990$, MSE = .346, $p = .013$,.

An examination of the ear advantages in Fig. 1 suggests that participants in the sad condition have both an attenuated REA in the emotional prosody condition, and an enhanced REA in the neutral condition. Separate Emotion Group (happy, angry, sad) × Ear ANOVAs were calculated for the neutral and emotional prosody conditions to further explore this relationship. Examining just the neutral trials, the Emotion Group × Ear interaction approached significance, $F(2, 93) = 2.316$, MSE = .32, $p = .105$, suggesting that participants in the sad group may have been more lateralised on the neutral trials than participants in the other two emotion groups. As participants were randomly assigned to condition,

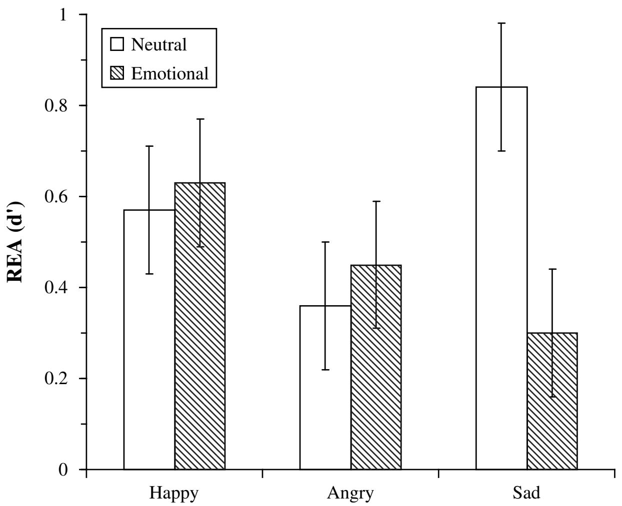


**Fig. 1.** Mean REAs in sensitivity, comparing emotional to neutral prosody for each emotion group. Error bars are standard errors of the mean difference between neutral and emotional conditions for each emotion group.

and there were no effects of block order that could explain a carry-over effect from the sad trials, this difference between the emotion groups in the neutral baseline condition most likely reflects chance variation between groups. For the emotion trials, the Emotion Group × Ear interaction was not significant, $F(2, 93) = 1.508$, MSE = .295, ns. However, this equivalence must be seen in light of the group differences on the neutral trials, keeping in mind that the comparison of neutral to emotion trials is made within-subjects, whereas the comparisons between different emotional prosodies is made between-subjects.

*3.2.2. Response time analyses*

Mean response times for hits were analyzed in a similar 2 (Prosody) × 2 (Ear) × 3 (Emotion Group) mixed model ANOVA. Responses were faster to neutral than to emotional prosodies, $F(1, 93) = 9.067$, MSE = 12,004, $p = .003$, and there was an overall REA, $F(1, 93) = 17.413$, MSE = 6371, $p < .001$. The Prosody × Ear interaction approached significance, $F(1, 93) = 2.724$, MSE = 2598, $p = .102$, reflecting an attenuated REA for the emotional prosodies ($M = 25$ ms, SD = 93 ms) relative to the neutral prosody ($M = 43$ ms, SD = 95 ms). Although the Prosody × Ear × Emotion Group interaction did not reach significance, $F(2, 93) = 1.835$ MSE = 2598, $p = .165$, further analyses were done for each emotion group separately to parallel the analysis done with the $d'$ values (see Fig. 2). For participants who heard the sad tone of voice, main effects of Prosody, $F(1, 31) = 4.126$, MSE = 15,566, $p = .05$, and Ear, $F(1, 31) = 7.542$, MSE = 6891, $p = .010$ were qualified by a Prosody × Ear interaction $F(1, 31) = 4.690$, MSE = 3323, $p = .038$. The REA for the emotional condition ($M = 18$ ms, SD = 68 ms) was attenuated relative to the REA for the neutral condition ($M = 62$ ms, SD = 126 ms). In contrast, participants in the happy condition demonstrated only a marginal effect of Ear, $F(1, 31) = 3.717$, MSE = 7384, $p = .063$, that did not interact with Prosody $F(1, 31) = .037$, ns. Similarly, participants in the angry prosody condition produced a significant REA, $F(1, 31) = 6.936$, MSE = 4849, $p = .013$, that did not interact with Prosody, $F(1, 31) = .487$, ns. In addition, participants in the angry prosody condition demonstrated an overall reaction time advantage for words spoken in a neutral voice Thus, while the Prosody × Ear × Emotion Group interaction did not reach significance in the response time data, the pattern of results is identical to that observed in the accuracy data. Specifically, participants who heard the sad prosody showed a significant attenuation of the REA in the emotion condition relative to the neutral condition, but those who heard happy or angry prosodies showed equivalent REAs for speech in the neutral and emotion conditions.

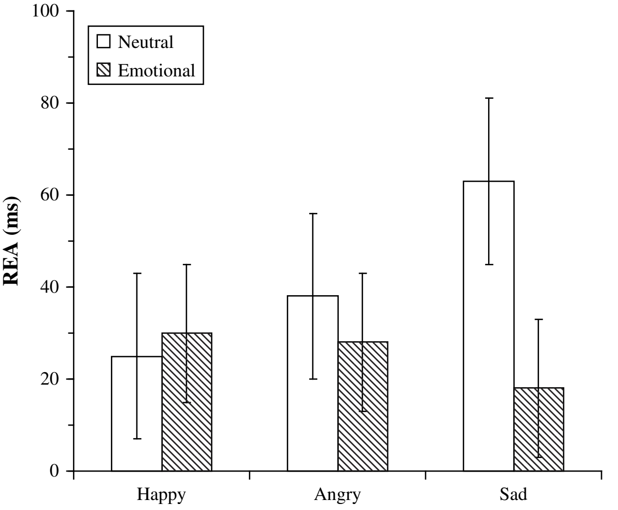


**Fig. 2.** Mean REAs in ms, comparing emotional to neutral prosody for each emotion group. Error bars are standard errors of the mean difference between neutral and emotional conditions for each emotion group.

Just as in the sensitivity data, examination of the response time data suggests that participants in the sad condition may have been both more lateralised on the neutral trials and less lateralised on the emotion trials than participants in the other emotion conditions. However, separate Emotion Group × Ear ANOVAs for neutral and emotional trials revealed no interaction on either neutral trials, $F(2, 93) = 1.133$, MSE = 4518, $p = .332$, or emotional trials, $F(2, 93) = .153$, MSE = 4451, $p = .858$. Again, the critical comparison is the within-subject comparison of neutral to emotional trials, for which participants in the sad condition show significant attenuation of the REA for emotion trials, and participants in the happy and angry conditions do not.

*3.2.3. Acoustic variables*

Given that an attenuation of the ear advantage was observed with sad prosody but not with angry or happy prosodies, a final analysis examined the acoustic parameters of the stimuli. One-way ANOVAs on each frequency parameter (treating the individual stimuli as subjects) revealed that the different tones of voice varied in mean $F_0$, $F(3, 12) = 18.772$, MSE = 424, $p < .001$, median $F_0$, $F(3, 12) = 19.526$, MSE = 435, $p < .001$ and $F_0$ standard deviation, $F(3, 12) = 49.008$, MSE = 71, $p < .001$. Post-hoc Tukey tests (with an adjusted alpha level of .05) revealed that, while happy and angry prosodies did not differ on any of the parameters, sad differed from the other emotional prosodies in all parameters. However, sad prosody did not differ from neutral in mean $F_0$ or median $F_0$, but only in $F_0$ standard deviation. For all parameters, sad prosody was *more* similar to neutral than were happy or angry prosodies, suggesting that the attenuation of the REA observed with sad prosody is unlikely to be driven by these acoustic differences.

A second acoustic analysis concerned the speech rate associated with the different prosodies. Poeppel (2003) has theorized that the right hemisphere samples over a window of 150–250 ms, as opposed to the left hemisphere sampling over windows of 20–40 ms. If sad speech is sufficiently slow, it may be amenable to right hemisphere phonological analysis. A one-way ANOVA on the ratings of duration revealed a marginal effect of Emotion, $F(3, 12) = 2.542$, MSE = 586, $p = .107$. Post-hoc Tukey tests revealed that none of the prosodies differed significantly from each other, but sad prosody (the slowest) was marginally slower than neutral (the fastest), $p = .102$. Given that this analysis was greatly underpowered, a tentative conclusion is that duration of the first phoneme is a possible acoustic variable that differentiates sad from neutral prosody. However, further exploration of the acoustic hypothesis will require a much larger sample of stimuli than were used in the present study.

## 4. General discussion

This study compared the effects of happy, angry, and sad emotional prosodies on hemispheric specialisation for a linguistic task. The findings from Experiment 1 confirmed that the stimuli used in the present study displayed the typical properties of a REA (reflecting left hemisphere specialisation) when participants attended to the linguistic dimension, and a LEA (reflecting right hemisphere specialisation) when participants attended to the prosodic dimension. Furthermore, valence effects were not observed; processing of happy, angry, and sad prosodies were equally lateralised to the right hemisphere.

In Experiment 2, participants performed the linguistic task while the emotional prosody of the voice was manipulated. Results for the sad prosody replicated those reported previously (Grimshaw et al., 2003). Specifically, the REA that was observed when the words were spoken in a neutral prosody was attenuated when the words were spoken in a sad prosody. As this replication involved different voice recordings, spoken by a different speaker in a different accent, the similarity of the results suggests that they are not due to idiosyncratic factors in the speaker's voice.

In contrast, happy and angry emotional prosodies had no effect on lateralisation for the linguistic task. This dissociation has some interesting implications for the interpretation of these results. Our initial hypothesis was that right hemisphere activation associated with emotional prosody facilitated linguistic processing. However, if the facilitation was driven by emotional prosody per se, then it should have been observed for all three prosodies, given that all three were lateralised equally to the right hemisphere in Experiment 1, and a preliminary ratings experiment found all three to be equivalent in emotional intensity. Therefore, we suggest that right hemisphere activation associated with the presence of emotional prosody is not sufficient to facilitate right hemisphere linguistic processing.

This interpretation is somewhat tentative, given that the sad prosody group appear to be more lateralized (although only in the sensitivity data) than the other two groups on the neutral trials. Thus it is possible that the influence of emotional prosody in this group reflects some pre-existing differences in hemispheric specialisation that were not controlled through the use of random assignment and matching of groups on age, handedness, and gender. We think this explanation is unlikely given that findings for the sad prosody group replicate those observed in Grimshaw et al. (2003). If one takes each group's performance on neutral trials as a baseline, then those listening to sad prosody showed attenuation of the REA, but those listening to other prosodies did not.

There are two alternative hypotheses that can explain the differences between the effects of the sad tone of voice and the other prosodies; the first in terms of psychoacoustic properties, and the second in terms of emotion. Acoustically, one of the defining features of sad prosody is that speech is produced at a relatively slow tempo (Johnstone & Scherer, 2000; Scherer, 1986). Poeppel and colleagues (e.g., Boemio, Fromm, Braun, & Poeppel, 2005; Poeppel, 2003) have argued that the left and right hemispheres' relative contributions to speech processing are driven by a difference in sampling rate. Analysis of the duration of the initial plosives of the stimuli in the present study revealed that the initial phonemes for sad stimuli were longer than those for neutral, happy, or angry stimuli. Given that the task could be completed on the basis of identification of the first phoneme, it is possible that the speech rate was sufficiently slowed to permit the right hemisphere, even with its slower sampling rate, to extract the necessary phonemic information. The acoustical analysis in this study was exploratory in nature, and a more thorough examination of the effects of speech rate would require a much larger stimulus set on which to compare the speech rate of different prosodies. This hypothesis could also be tested more directly by manipulating the speech rate of the stimuli. Future research in this area should explore the relative contributions of acoustic variables to the effects of emotional prosody on linguistic processing.

Alternatively, the results may be driven by the emotional properties of sad prosody. Although the right hemisphere is specialised for the comprehension of all emotions, there are emotion-specific differences in the lateralisation of emotional experience. The findings from the present study are most consistent with the theorizing of Davidson and colleagues who associate right frontal activation with withdrawal-motivated emotions such as sadness and fear, and left frontal activation with approach-motivated emotions such as happiness and anger (Davidson et al., 1990; Harmon-Jones & Sigelman, 2001; Harmon-Jones, Vaughn-Scott, Mohr, Sigelman, & Harmon-Jones, 2004). In fact, if the effects of emotional prosody are somehow additive with those of linguistic processing, one might expect an enhancement of the REA for words spoken with happy or angry prosody, in contrast to the results observed here. Perhaps right frontal activation associated with sad experience facilitated right hemisphere linguistic processing, whereas left frontal activation associated with anger or happiness had no effect on linguistic processes that were already lateralised to the left hemisphere (cf. Asbjörnsen et al., 1992).

Electrophysiological or neuroimaging procedures may be better suited to examine the specific neural correlates associated with the perception of different emotional prosodies than the dichotic listening methodology used in the present study. Partial support for the emotion-specific hypothesis put forward here comes from the neuroimaging study of Buchanan et al. (2000) who used the same happy and sad stimulus recordings as were used in Grimshaw et al. (2003). In an fMRI study, they found that attending to both the happy and sad prosodies (using attention to the linguistic dimension as a baseline) produced activation in right anterior auditory cortex, right inferior frontal and inferior parietal gyri, and in left cingulate. However, when the sad and happy prosodies were contrasted, activation in right anterior middle frontal gyrus was found to be unique to sad prosody, whereas no brain areas were found to be unique to happy prosody. Thus, even though comprehension of both happy and sad prosody are equally lateralised to the right hemisphere as assessed by dichotic listening methodology, functional imaging reveals right frontal activity that is unique to perception of sad prosody. Buchanan et al. did not include angry prosody in their study, and so direct comparisons of angry and sad are not possible. However, Grandjean et al. (2005) found that angry voices (relative to neutral voices) elicited activity in bilateral (predominantly right) superior temporal sulci, even for a gender identification task for which the emotional prosody was irrelevant. Task differences make comparisons to the Buchanan et al. study difficult however, and any strong conclusions about the functional anatomy of different prosodies awaits a study in which they are all directly compared using the same task. The growing neuroimaging literature in this area serves to underscore the fact that, while valence effects for prosodic perception are not typically observed in dichotic listening or patient studies (cf., Erhan et al., 1998), these methodologies may underestimate neurological distinctions between different emotions.

Although the present study cannot precisely pinpoint the mechanism by which sad prosody facilitates right hemisphere contributions to linguistic processing, it does demonstrate that this facilitation is robust, and that it cannot simply be mediated by activation of right temporal structures associated with prosodic comprehension. Emotional prosodies differ on both psychoacoustic and affective dimensions, both of which are associated with hemispheric differences in processing. Future research in this area must therefore target these contrasting theories more specifically. For example, fearful voices are similar to sad voices in withdrawal motivation, but are more similar to happy and angry voices in speech rate (Scherer, 1986). If right hemisphere linguistic processing is mediated by withdrawal-associated motivation, then an attenuated REA should be observed for fearful prosody; if it is driven by a slow speech rate then fearful prosody should not show this effect.

Regardless of the specific mechanism by which sad prosody alters the neurological correlates of linguistic processing, this finding raises some interesting questions regarding the consequences of a shift toward right hemisphere processing of sad speech. Right hemisphere linguistic capabilities are known to differ from those of the left in qualitative ways. For example, the right hemisphere is more sensitive to emotional than non-emotional words (Atchley, Stringer, Mathias, Ilardi, & Minatrea, 2007), and is more likely to activate both dominant and subordinate meanings of ambiguous words (Burgess & Simpson, 1988; Koivisto, 1997; Titone, 1998). These different patterns of semantic activation are likely associated with the right hemisphere's important contributions to metaphor processing (Bottini et al., 1994; Mashal, Faust, Hendler, & Jung-Beeman, 2007; Schmidt, DeBuse, & Seger, 2007), humour (Coulson & Williams, 2005) and discourse processing (Beeman, 1998; Brownell, Potter, Bihrle, & Gardner, 1986). Almost all research on hemispheric differences in semantic organisation has relied on visual word presentation, and not speech. Thus, whether emotional prosody affects hemispheric patterns of semantic activation remains to be determined. The present study illustrates the value of considering both linguistic and prosodic dimensions together when developing neurolinguistic models of language processing.

## Acknowledgements

We wish to thank Matthew Duignan for his vocal and technical talents, Sharon Fisher for library assistance, and Laurel O'Brien for her assistance in analysis of the sound files.

## References

Asbjörnsen, A., Hugdahl, K., & Bryden, M. P. (1992). Manipulations of subjects' level of arousal in dichotic listening. *Brain and Cognition, 19*, 183–194.

Atchley, R. A., Stringer, R., Mathias, E., Ilardi, S. S., & Minatrea, A. D. (2007). The right hemisphere's contribution to emotional word processing in currently depressed, remitted depressed, and never-depressed. *Journal of Neurolinguistics, 20*(2), 145–160.

Beeman, M. (1998). Coarse semantic coding and discourse comprehension. In M. Beeman, & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 255–284). Mahwah, NJ: Erlbaum.

Beeman, M. J., & Chiarello, C. (1998). Complementary right- and left-hemisphere language comprehension. *Current Directions in Psychological Science, 7*(1), 2–8.

Berckmoes, C., & Vingerhoets, G. (2004). Neural foundations of emotional speech processing. *Current Directions in Psychological Science, 13*(5), 182–185.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience, 8*(3), 389–395.

Boersma, P., & Weenink, D. (2007). *Praat: Doing phonetics by computer* (Version 4.6.06).

Borod, J. C., Cicero, B. A., Obler, L. K., Welkowitz, J., Erhan, H. M., Santschi, C., et al. (1998). Right hemisphere emotional perception: evidence across multiple channels. *Neuropsychology, 12*, 446–458.

Borod, J. C., Zgaljardic, D., Tabert, M. H., & Koff, E. (2001). Asymmetries of emotional perception and expression in normal adults. In G. Gainotti (Ed.), *Handbook of neuropsychology: Emotional behaviour and its disorders*. Oxford UK: Elsevier.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: a positron emission tomography activation study. *Brain: A Journal of Neurology, 117*(6), 1241–1253.
Bowers, D., Coslett, H. B., Bauer, R. M., Speedie, L. J., & Heilman, K. M. (1987). Comprehension of emotional prosody following unilateral hemispheric lesions: processing defect versus distraction defect. *Neuropsychologia, 25*(2), 317–328.
Brownell, H. H., Potter, H. H., Bihrle, A. M., & Gardner, H. (1986). Inference deficits in right brain-damaged patients. *Brain and Language, 27*(2), 310–321.
Bryden, M. P. (1988). An overview of the dichotic listening procedure and its relation to cerebral organization. In K. Hugdahl (Ed.), *Handbook of dichotic listening: Theory, methods, and research* (pp. 1–44). Chichester: John Wiley and Sons.
Bryden, M. P., & MacRae, L. (1989). Dichotic laterality effects obtained with emotional words. *Neuropsychiatry, Neuropsychology and Behavioural Neurology, 1*, 171–176.
Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., et al. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research, 9*, 227–238.
Burgess, C., & Simpson, G. (1988). Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language, 33*(1), 86–103.
Charbonneau, S., Scherzer, B. P., Aspirot, D., & Cohen, H. (2003). Perception and production of facial prosodic emotions by chronic CVA patients. *Neuropsychologia, 41*(5), 605–613.
Coulson, S., & Williams, R. F. (2005). Hemispheric asymmetries and joke comprehension. *Neuropsychologia, 43*(1), 128–141.
Davidson, R. J. (1992). Emotion and affective style: hemispheric substrates. *Psychological Science, 3*(1), 39–43.
Davidson, R. J. (1993). Parsing affective space: perspectives from neuropsychology and psychophysiology. *Neuropsychology, 7*(4), 464–475.
Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology: I. *Journal of Personality and Social Psychology, 58*(2), 330–341.
Elias, L. J., Bryden, M. P., & Bulman-Fleming, M. B. (1998). Footedness is a better predictor than is handedness of emotional lateralisation. *Neuropsychologia, 36*, 37–43.
Erhan, H., Borod, J. C., Tenke, C. E., & Bruder, G. E. (1998). Identification of emotion in a dichotic listening task: event-related brain potential and behavioral findings. *Brain and Cognition, 37*, 286–307.
Friederici, A. D., & Alter, K. (2004). Lateralisation of auditory language functions: a dynamic dual pathway model. *Brain and Language, 89*(2), 267–276.
Gadea, M., Gómez, C., González-Bono, E., Espert, R., & Salvador, A. (2005). Increased cortisol and decreased right ear advantage (REA) in dichotic listening following a negative mood induction. *Psychoneuroendocrinology, 30*, 129–138.
Gainotti, G., Caltagirone, C. G. M., & Masullo, C. (1981). Selective semantic-lexical impairment of language comprehension in right-brain-damaged patients. *Brain and Language, 13*, 201–211.
Gandour, J., Wong, D., Dzemidzic, M., Lowe, M., Tong, Y., & Li, X. (2003). A Cross-linguistic fMRI study of perception of intonation and emotion in Chinese. *Human Brain Mapping, 18*(3), 149–157.
Gazzaniga, M. S., & Sperry, R. W. (1967). Language after section of the cerebral hemispheres. *Brain, 90*, 131–148.
Gold, B. T., & Kertesz, A. (2000). Right hemisphere semantic processing of visual words in an aphasic patient: an fMRI study. *Brain and Language, 73*(3), 456–465.
Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience, 8*(2), 145–146.
Grimshaw, G. M. (1998). Integration and interference in the cerebral hemispheres: relations with hemispheric specialisation. *Brain and Cognition, 36*(2), 108–127.
Grimshaw, G. M., Kwasny, K. M., Covell, E., & Johnson, R. A. (2003). The dynamic nature of language lateralisation: effects of lexical and prosodic factors. *Neuropsychologia, 41*(8), 1008–1019.
Harmon-Jones, E., & Sigelman, J. (2001). State anger and prefrontal brain activity: evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology, 80*(5), 797–803.
Harmon-Jones, E., Vaughn-Scott, K., Mohr, S., Sigelman, J., & Harmon-Jones, C. (2004). The effect of manipulated sympathy and anger on left and right frontal cortical activity. *Emotion, 4*(1), 95–101.
Heller, W. (1990). The neuropsychology of emotion: developmental patterns and implications for psychopathology. In N. L. L. Stein, Bennett., & Tom Trabasso (Eds.), *Psychological and biological approaches to emotion* (pp. 167–211). Hillsdale, NJ/ England: Lawrence Erlbaum Associates.
Heller, W. (1993). Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychology, 7*, 476–489.
Johnson, K. (1997). *Acoustic and auditory phonetics*. Cambridge: Blackwell.
Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis, & J. Haviland (Eds.), *Handbook of emotion* (2nd ed.). (pp. 220–235) New York: Guilford.
Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology, 15*, 166–171.
Koivisto, M. (1997). Time course of semantic activation in the cerebral hemispheres. *Neuropsychologia, 35*(4), 497–504.
Kucharska-Pietura, K., Phillips, M. L., Gernand, W., & David, A. S. (2003). Perception of emotions from faces and voices following unilateral brain damage. *Neuropsychologia, 41*(8), 1082–1090.
Lattner, S., Meyer, M. E., & Friederici, A. D. (2005). Voice perception: sex, pitch, and the right hemisphere. *Human Brain Mapping, 24*(1), 11–20.
Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2007). An fMRI investigation of neural correlates underlying the processing of novel metaphoric expressions. *Brain and Language, 100*(2), 115–126.
Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia, 41*(10), 1410–1421.
Pell, M. D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language, 96*, 221–234.
Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralisation as 'asymmetric sampling in time'. *Speech Communication, 41*, 245–255.

Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychological Bulletin, 99*(2), 143–165.
Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Communication, 40*(1–2), 227–256.
Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*(1), 24–30.
Schmidt, G. L., DeBuse, C. J., & Seger, C. A. (2007). Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. *Brain and Language, 100*(2), 127–141.
Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime suite*. (Version 1.0). Pittsburgh: Psychology Software Tools Inc..
Stirling, J., Cavill, J., & Wilkinson, A. (2000). Dichotically presented emotionally intoned words produce laterality differences as a function of localisation task. *Laterality, 5*, 363–371.
Titone, D. (1998). Hemispheric differences in context sensitivity during lexical ambiguity resolution. *Brain and Language, 65*(3), 361–394.
Van Lancker, D., & Sidtis, J. J. (1992). The identification of affective-prosodic stimuli by left- and right-hemisphere damaged subjects: all errors are not created equal. *Journal of Speech and Hearing Research, 35*, 963–970.
Van Lieshout, P. (2004). PRAAT: short tutorial. Available from: http://www.slp.utoronto.ca/Assets/Misc+Resarch+Lab+Files/assets/PrattMan.pdf. Accessed 20.10.08.
Van Strien, J. W., & Boon, C. A. (1997). Altered visual field asymmetry for lexical decision as a result of concurrent presentation of music fragments of different emotional valences. *Journal of the International Neuropsychological Society, 3*, 473–479.
Walker, J. P., Daigle, T., & Buzzard, M. (2002). Hemispheric specialisation in processing prosodic structures: revisited. *Aphasiology, 16*(12), 1155–1172.
Zaidel, E. (1976). Auditory vocabulary of the right hemisphere following brain bisection or hemidecortication. *Cortex, 12*, 191–211.