

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/30209> holds various files of this Leiden University dissertation

Author: Witteman, Juriaan

Title: Towards a cognitive neuroscience of prosody perception and its modulation by alexithymia

Issue Date: 2014-12-18

Chapter 2.2

The nature of hemispheric specialization for prosody perception: ERP evidence

Witteman, J., Goerlich, K. S., Martens, S., Aleman, A., Van Heuven, V. J., & Schiller, N. O. (2014). The nature of hemispheric specialization for prosody perception. *Cognitive Affective & Behavioral Neuroscience, 14*, 1104-1114.

Abstract

Recent evidence suggests that there is relative right hemispheric specialization for emotional prosody perception while linguistic prosody perception is under bilateral control. It is still unknown, however, how hemispheric specialization for prosody perception might arise. Two main hypotheses have been put forward. Cue-dependent hypotheses on the one hand, propose that hemispheric specialization is driven by specialization for non-prosody specific processing of acoustic cues. The functional lateralization hypothesis on the other hand, proposes that hemispheric specialization is dependent on the communicative function of prosody with emotional and linguistic prosody processing lateralized to the right and left hemisphere, respectively. In the present study, the functional lateralization hypothesis of prosody perception was systematically tested by instructing one group of participants to evaluate the emotional prosody and another group to the linguistic prosody dimension of bi-dimensional prosodic stimuli in a dichotic listening paradigm while event related-potentials (ERPs) were recorded. The results showed that the right ear advantage was associated with a decreased latency of an early negativity in the contralateral hemisphere. No evidence was found for functional lateralization. These findings suggest that functional lateralization effects for prosody perception are small and support the structural model of dichotic listening.

1. Introduction

Speech prosody is an important means to convey emotions (e.g., indicating whether the speaker is angry or sad) and linguistic structure (e.g., signifying a statement or a question). A longstanding question in the neuropsychological literature has been whether there is hemispheric specialization for the perception of prosody. Recent meta-analytic evidence suggests that there is a *relative* right hemispheric specialization for emotional prosody perception, while the processing of linguistic prosody seems to be controlled bilaterally (Witteman, Van IJzendoorn, Van de Velde, Van Heuven, & Schiller, 2011).

However, the nature of this relative right hemispheric specialization for emotional prosody is currently unknown. Two mechanisms have been proposed as to *how* hemispheric specialization for prosody perception might arise. On the one hand, *cue-dependent* lateralization hypotheses propose that right-hemisphere specialization for emotional prosody perception can be explained by a (non-prosody specific) advantage of the right hemisphere for early acoustic processing, such as spectral processing (Van Lancker & Sidtis, 1992), as spectral parameters appear to be particularly important for decoding emotional prosody (Scherer, 2003). Indeed, a recent meta-analysis of the neuroimaging literature of emotional prosody perception found preliminary evidence for relative right hemispheric specialization of the primary and secondary auditory cortex (Witteman, Van Heuven, & Schiller, 2012), which could be interpreted as (indirect) support for the cue-dependent lateralization hypothesis.

On the other hand, the *functional* lateralization hypothesis posits that the right hemisphere is specialized in the processing of the emotional communicative function of emotional prosody (Van Lancker, 1980), whereas the left hemisphere might be specialized in the processing of linguistic prosodic function. The dynamic dual pathway model by Friederici and Alter (2004) further suggests that when linguistic prosody is more bound to segments (such as in the case of metrical stress), it is left lateralized, while linguistic prosody at the sentence level (such as boundary marking) is right lateralized. The goal of the present investigation was to systematically test whether there is functional lateralization for prosody perception.

To comprehensively test the functional lateralization hypothesis it is necessary to vary the communicative function of prosodic material while keeping acoustics constant, and to observe whether there is a change in the difference of activity between the hemispheres. Note that the acoustic and functional lateralization hypotheses are non-mutually exclusive and could represent different stages of the prosody perception process. Indeed, recent neural models of prosody perception have suggested that there might be acoustically driven lateralization in an initial processing stage, but more semantically (functionally) driven lateralization in subsequent stages (Bruck, Kreifelts, & Wildgruber, 2011; Kotz & Paulmann, 2011). To shed light on the issue of when in time functional lateralization arises the present ERP study systematically manipulated the function of prosody by instructing one group of subjects to evaluate the emotional prosody dimension and a different group of participants to evaluate the linguistic prosody dimension of identical bi-dimensional stimuli.

Interestingly, Paulmann, Jessen and Kotz (2012) recently reported such a direct comparison of emotional to linguistic prosody perception using the cross-splicing paradigm. A so called 'prosodic expectancy positivity (PEP)' was found that was more pronounced for emotional than linguistic prosody expectancy violations, suggesting prioritized processing of emotional prosodic cues. However, the authors did not find

task-driven effects and suggested that the absence of such effects at the electrophysiological level might have been caused by a lack of statistical power to detect (presumably small) task effects, requiring investigation of this issue using larger study samples. Furthermore, a within-subjects manipulation of task demands was employed, which required subjects to switch between an emotional and a linguistic task set. As the authors pointed out, this procedure might have reduced task-driven differences in prosody processing, warranting a more extended investigation of this issue using a between subjects manipulation of task demands.

The dichotic listening (DL) paradigm is particularly suited to study hemispheric specialization in the auditory modality (Greenwald & Jerger, 2003). In the divided-attention version of this paradigm, two different stimuli are presented to each ear. The participant has to divide attention over the auditory channels and react only to the target stimulus presented to one of the ears. It is generally agreed that in the DL-situation the ipsilateral projection of information from the ears to the cerebral hemispheres is inhibited, rendering auditory information from the ears to be primarily available to the contralateral hemisphere (e.g. see Davidson & Hugdahl, 1995). Therefore, if there were hemispheric specialization for the processing of an auditory stimulus, this would be observable as a performance advantage of the ear contralateral to the specialized hemisphere as it has direct access to the specialized processing module (for a discussion of the exact mechanisms behind DL, see Grimshaw, Kwasny, Covell, & Johnson, 2003).

Only one previous study has combined the DL-paradigm with ERPs to study lateralization of emotional prosody perception. Erhan, Borod, Tenke and Bruder (1998) presented participants with dichotic pairs of nonsense syllables, each of which was spoken in one of seven emotional prosodic categories. Participants had to detect a pre-specified emotional prosodic category as fast as possible while the EEG was recorded. Auditory target detection studies like these generally find an initial negativity followed by a positivity for targets as compared with non-targets (e.g., see Fitzgerald & Picton, 1983). In the case of the positivity, it has been demonstrated that its amplitude increases when target probability decreases, while for the earlier negativity this effect is weak at best (Polich & Bondurant, 1996). The positivity can be subdivided in the P3b component and the slow wave, which together have been termed Late Positive Potential or LPP (e.g. see Briggs & Martin, 2009). Erhan et al. (1998) indeed found an N1 and a sustained negativity, followed by a late positivity and a slow wave. Further, at the behavioral level a left ear advantage (LEA) was found for accuracy, in line with a right-hemispheric specialization for emotional prosody. The sustained negativity (300-879 ms post stimulus onset) was identified as a potential electrophysiological marker of the behavioral ear advantage, hypothesized to reflect the emotional categorization process. However, although the relatively late latency of the component might be interpreted to reflect fairly late (and therefore possibly functional) processing, strictly speaking it is unclear whether this lateralized component reflected early acoustic lateralization, more abstract functional hemispheric specialization or both as the function of prosody was not manipulated independently of the acoustics (or vice versa).

In the present ERP study, linguistic and emotional prosodic task demands were manipulated between-subjects with comparably high statistical power while keeping acoustics constant in a divided-attention dichotic auditory target detection. We predicted that, if there is functional lateralization of prosodic perception, there should be a right-ear advantage for the linguistic prosody perception task and a left-ear

advantage for the emotional task. At the electrophysiological level, we hypothesized that the distribution of the N2 or LPP over the hemispheres might change correspondingly, demonstrating functional lateralization. Lastly, we considered the possibility that an earlier component (reflecting acoustic processing) might also be sensitive to variation in task demands, reflecting top-down task dependent (and possibly lateralized) modulation of earlier acoustic processing, as has been demonstrated previously (Sussman, Winkler, Huotilainen, Ritter, & Näätänen, 2002).

2. Methods

2.1. Participants

A total of 82 subjects participated in the experiment. Five subjects were excluded because of lower than chance level performance and three due to noisy EEG-data, resulting in a total of 74 participants; 41 (21 male; mean age 23, range 18-37) for the emotional prosodic task and 33 different participants (16 male; mean age 23, range 19-36) for the linguistic prosodic task. The two groups did not differ in male-female ratio or age (for all: one-way ANOVA, $F < 1.01$, $p > 0.5$) All participants were right-handed as assessed with the Edinburgh Handedness Inventory (Oldfield, 1971), were native speakers of Dutch, had normal or corrected-to-normal vision and no (neuro)psychiatric problems in present or past. Participants showing a mean inter-ear hearing threshold difference greater than 12 dB on .5 kHz, 1 kHz, 2 kHz, 4kHz sinusoid tones were excluded from the study. Participants received 20 Euros for their participation in the two hour EEG session. Informed consent was obtained from all participants prior to the experiment. The study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki.

2.2. Materials

Pseudowords (see Appendix A) with a bisyllabic structure were generated by randomly combining monosyllables that were composed of random combinations of an initial consonant, a vowel and a final consonant. All pseudowords obeyed Dutch phonotactics and were checked to verify the absence of semantic content. All pseudowords were uttered with angry and sad prosody and with stress on the first and second syllable by a professional actress and recorded at 16 bit resolution and a 44.1 kHz bit sampling rate in a sound-proofed booth. Items were intensity normalized and had a mean duration of 750 msec. In line with previous literature (Cutler, 2005; Scherer, 2003) unstressed syllables differed from stressed syllables primarily in duration while sad and angry prosody additionally differed in F_0 , F_0 variability, and variation in intensity (see Table 1). Note that angry and sad items did not show a large intensity difference as the stimuli had been intensity-normalized. To verify the validity of the intended prosodic contrasts, a panel of five healthy volunteers classified each sad and angry prosodic stimulus (in addition to neutral, happy and surprised prosodic stimuli that were not used in the present study) in a forced choice task and rated each item on a five-point typicality scale (1 = very atypical, 5 = very typical). Only pseudowords were selected for which the emotional prosodic contrasts (angry and sad intonation) were classified correctly by at least 4 out of 5 panel members and which had a typicality rating of at least 3.5.

There is evidence that increasing spectral overlap between the target and the competing stimulus increases the suppression of ipsilateral afferent routes from the ears to the cerebral hemispheres, which enhances ear advantages (Della Penna et al., 2007).

Therefore, to maximize spectral overlap, a competing babble stimulus was created by selecting random small segments of speech of the actress and superimposing them onto each other. Dichotic stimuli were created by selecting a random sample of the babble stimulus with the exact same duration of the target stimulus and combining the two (with the target presented in one channel and the babble stimulus in the other channel). This procedure ensured constant and high competition between the target stimulus and the distracter.

Two dichotic target detection tasks were created that presented identical prosodic dichotic targets. From the pool of validated dichotic stimuli, six that had sad prosody and stress on the first syllable served as targets. For the emotional task, subjects were instructed to press as fast as possible when they heard a sad target. For the linguistic task, subjects responded as fast as possible when targets had word initial stress. Hence, identical stimuli served as targets for both tasks while only the task demands (emotional versus linguistic decision) varied. Ten items with angry prosody and stress on the second syllable served as non-targets in both target detection tasks. Additionally, four task-specific (non-target) filler items were presented ten times to each ear but were not analyzed. For the emotional task, angry items with stress on the first syllable were added to prevent subjects from using a linguistic strategy to detect emotional prosody. For the linguistic task, sad items with stress on the second syllable were added to prevent subjects from using an emotional strategy to detect stress position. Each item was presented ten times to each ear, resulting in a total of 120 target and 280 non-target trials per ear. Hence, a target occurred in 30% of the trials for both the emotional and the linguistic DL-task, and target probability was 50% between the two ears. Task-irrelevant prosodic categories (e.g. word-initial versus word-final stress for the emotional task) had a 50% probability of occurring.

Table 1. Acoustic properties of the linguistic and emotional prosodic contrasts (SD in parentheses).

	Stressed syllable	Unstressed syllable	Angry word	Sad word
Mean intensity (dB)	74.22 (2.56)	73.28 (2.41)	72.51 (1.18)	73.94 (1.57)
Mean F₀ (Hz)	239.62 (54.71)	219.95 (36.59)	255.30 (13.81)	202.37 (35.35)
Total duration (s)	0.37 (0.058)	0.30 (0.05)	0.89 (0.057)	0.71 (0.018)
Mean variation (SD) F₀	57.53 (34.29)	31.98 (15.30)	65.88 (9.88)	42.66 (16.47)
Mean variation (SD) intensity	8.70 (2.74)	6.07 (1.65)	13.68 (1.78)	6.67 (1.40)

2.3. Procedure

Each participant completed one of the two dichotic target-detection tasks in a sound-proof booth. Subjects were instructed that they would hear a prosodic stimulus in one ear and people babbling in the other ear. They were told to ignore the babble stimulus and to decide as fast and accurately as possible when they heard the target prosody by pressing the spacebar with the index finger. Response hand was counterbalanced across participants. Participants were instructed that they could respond while the stimulus was still playing (i.e., RTs were recorded from the onset of the stimulus). All instructions were exactly the same for both tasks except for the description of the task-specific prosodic stimulus categories.

Each dichotic listening task started with twelve dichotic practice trials. Participants kept practising these items until a performance level of at least 75% correct was reached. Subsequently, the experimental trials started which encompassed a total of 400 dichotic trials (120 targets, 200 non-targets and 80 fillers). Stimulus presentation order was (pseudo-) random with the restriction that no more than two consecutive presentations of a target were allowed.

An experimental trial started with a black fixation cross that was presented for 1500 ms. Subsequently, a red fixation cross and a binaural warning tone of 500 Hz were presented for 500 ms, after which the dichotic stimulus was presented while the fixation cross remained red. The trial ended 2000 ms after stimulus onset or when participants made a response. Participants were instructed to fix their gaze on the fixation cross throughout the experiment in order to reduce eye movements and not to blink while the fixation cross was red (i.e., when the stimuli were presented).

Stimulus presentation was controlled using E-prime 1.2 and stimulus material was presented at 16 bit resolution and a 44.1 kHz sampling frequency through headphones at a comfortable intensity level.

2.4. EEG Recordings

Electroencephalogram (EEG) was recorded from 64 tin electrodes mounted in an elastic electro-cap organized according to the international 10/20 system. EEG data were recorded with a linked mastoid physical reference and were re-referenced using an average reference. Vertical and horizontal electrooculogram (VEOG and HEOG) were recorded with two pairs of electrodes, one pair placed above and below the left eye, the other pair placed beside the two eyes. The ground electrode was applied to the sternum. Impedance of all electrodes was kept below 5 k Ω for each participant. EEG was continuously recorded with a sampling rate of 500 Hz, amplified, and off-line digitally low-pass filtered with a cut-off frequency of 30 Hz. Participants were seated in front of a monitor at a distance of approximately 50 cm in a dimly lit, electrically shielded and sound-attenuated booth.

2.5. Behavioral analysis

A RM-MANOVA was performed with Ear (left, right) as a within-subject factor, and Task (emotional, linguistic) as between-subject factor. Dependent measures were proportion correct responses and mean reaction time for correct responses.

2.6. ERP-analysis

The EEG data were analyzed with Brain Vision Analyzer (version 1.05). Prior to averaging, trials with eye-movement and blink artefacts were excluded from analysis. Criteria for artefact rejection within an epoch were a maximal voltage step of 50 μV , a maximal difference between two values in a segment of 100 μV , and a minimal and maximal amplitude of -100 μV and 100 μV , respectively.

ERP epochs were computed in a 1000 ms time-window and aligned to a 100 ms pre-stimulus baseline. Individual averages included at least 38 correctly detected target trials per ear. Based on the previous literature on attentive auditory target detection (e.g., see Fitzgerald & Picton, 1983; Nager et al., 2001; Becker & Reinvang, 2007), we expected to find negativities in the N2 range, followed by a late (> 300 ms) positive potential (LPP) including a P3b and a slow wave, when comparing targets to non-targets. Visual inspection of the ERP waves confirmed our expectations and additionally suggested an early negativity in the N1 latency range for targets as compared to non-targets (see Figure 2). Although both peak and mean amplitude analyses were considered, we decided to exclusively report mean amplitude analyses as only the early negativity exhibited a clear peak (as can be observed in Figure 2). For each participant, mean amplitudes for correct target detection were computed for 100-140 ms (early negativity), 180-320 ms (N2) and 350-900 ms (LPP). Because the early negativity exhibited a clear peak, peak latency could be analyzed additionally.

Due to the relatively large sample size of the present investigation, an analysis of the electrophysiological correlate of the ear advantage was possible. Participants were divided into a left ear advantage (LEA) group and a right ear advantage (REA) group, defined as the negative or positive difference between reaction times for correct responses of the left and right ear, respectively. This group analysis served to identify differences in ERP components depending on the laterality of the ear advantage (LEA versus REA).

For all analyses, a set of 36 electrodes was used (see Figure 1). For each hemisphere, the 18 electrodes were divided into six regions of interest (ROIs) comprising three electrodes each. Topographic effects were then analysed using two factors: Laterality (left hemisphere, proximal electrodes: F3, FC3, FC1, C3, C1, FC3, P3, PO3, P1; lateral electrodes: F7, F5, FC5, T7, C5, CP5, P7, P5, PO7; right hemisphere, proximal electrodes: F4, FC4, FC2, C4, CP4, C2, P4, PO4, P2; lateral electrodes: F8, F6, FC6, T8, C6, CP6, P8, P6, PO8), and a factor Longitudinality (left hemisphere, frontal electrodes: F7, F5, FC5, F3, FC3, FC1; central electrodes: T7, C5, CP5, C3, C1, CP3; posterior electrodes: P7, P5, PO7, P3, P1, PO3; right hemisphere, frontal electrodes: F8, F6, FC6, F4, FC4, FC2; central electrodes: T8, C6, CP8, C4, C2, CP4; posterior electrodes: P8, P6, PO8, P4, P2, PO4).

For the main analyses, a multivariate analysis of variance (MANOVA) was conducted using PASW Statistics with Ear (left, right), Hemisphere (left, right), Laterality (proximal, lateral) and Longitudinality (frontal, central, posterior) as within-subject factors and Task (linguistic, emotional) as between-subject factor and mean amplitude as dependent variable (for the N1 component latency was additionally analyzed as dependent variable). An additional MANOVA was performed using exactly the same factors but with the additional factor EA Group to investigate the electrophysiological correlates of the behavioral ear advantage. Greenhouse-Geisser corrected p -values are reported.

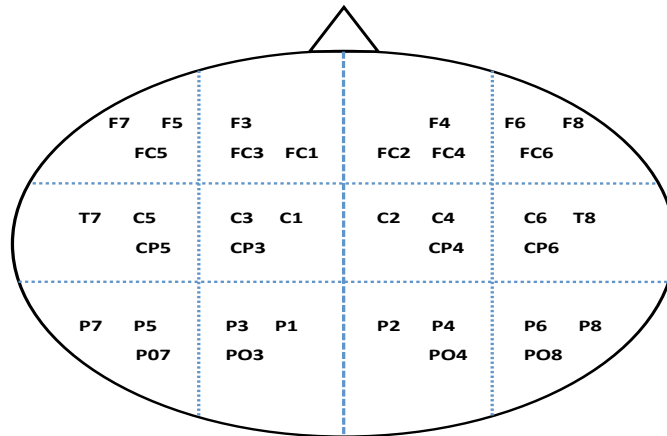


Figure 1. Electrode array used for the experiment. Dashed lines indicate the combination of the levels of the factor Laterality (central, lateral) and Longitudinality (anterior, central, posterior) for each Hemisphere (left, right) .

3. Results

3.1 Behavioral results

The behavioral results can be found in Table 2. Repeated Measures (RM)-MANOVA suggested that the emotional task was easier than the linguistic task as indicated by a main effect of Task ($F(2,71) = 91.54, p < 0.001, \eta_p^2 = .72$). Follow-up univariate tests confirmed that both reaction time was faster ($F(1,71) = 180.91, p < 0.001, \eta_p^2 = .72$) and accuracy higher ($F(1,71) = 57.84, p < 0.001, \eta_p^2 = .45$) for the emotional as compared to the linguistic task.

In sum, the behavioral results suggest that the emotional prosody task was easier than the linguistic prosody task, but no statistically reliable ear advantage was observed.

Table 2. Behavioral results for the experimental conditions (SD in parentheses).

	Linguistic task		Emotional task	
	Left ear	Right ear	Left ear	Right ear
Accuracy (% correct)	88 (0.10)	86 (0.11)	99 (0.02)	99 (0.01)
Reaction time (ms)	1016 (178)	1025 (191)	591 (83)	598 (76)

3.2. *Electrophysiological results –task effects*

The ERP waves of the early negativity, N2 and LPP are visualized in Figure 2. As the present paper is focused on task effects, in the results section below, only task-related effects and the most complete interactions are discussed. For the qualified main effects and simpler interactions and non-task related effects, we refer the reader to the Supplementary information.

3.2.1. *Early negativity*

The RM-MANOVA for mean amplitude revealed a Hemisphere \times Ear \times Longitudinality \times Task interaction ($F(2,142) = 3.34, p < 0.05, \eta_p^2 = .05$). Following up with separate ANOVAs for each level of Longitudinality, revealed that at frontal sites, the left ear elicited a larger negativity than the right for the emotional task in both hemispheres while for the linguistic task each ear elicited a larger negativity in the contralateral hemisphere ($F(1,72) = 6.25, p < 0.05, \eta_p^2 = .08$) while at central ($F(1,72) = 0.2, p > 0.05, \eta_p^2 = .003$) and posterior sites ($F(1,72) = 0.1, p > 0.05, \eta_p^2 = .001$) there was no three-way interaction.

An identical MANOVA with latency as dependent variable revealed a Hemisphere \times Ear \times Task interaction ($F(1,71) = 5.93, p < 0.05, \eta_p^2 = .08$) with shorter latency in the contralateral hemisphere for each ear for the emotional task, while for the linguistic task latency was shorter in the hemisphere ipsilateral to each ear.

In sum, at frontal sites the left ear elicited a stronger negativity than the right ear in both hemispheres for the emotional task, while for the linguistic task each ear elicited a larger negativity in the contralateral hemisphere. Last, latency was shorter in the hemisphere contralateral to the stimulated ear for the emotional task, while for the linguistic task latency was shorter in the ipsilateral hemisphere.

3.2.2. *N2*

The MANOVA for mean amplitude revealed a Longitudinality \times Task interaction ($F(2,142) = 19.74, p < 0.001, \eta_p^2 = .22$). Following up with separate ANOVAs for each level of Longitudinality revealed a larger N2 amplitude for the emotional task compared to the linguistic task at frontal sites ($F(1,72) = 13.24, p = 0.001, \eta_p^2 = .16$), a trend for the emotional task to elicit a greater negativity than the linguistic task at central sites ($F(1,72) = 3.76, p = 0.057, \eta_p^2 = .05$), and a greater negativity for the linguistic task than the emotional task at posterior sites ($F(1,72) = 6.04, p < 0.05, \eta_p^2 = .08$).

In sum, the emotional task elicited a larger negativity than the linguistic task at frontocentral sites, while the linguistic task elicited a larger negativity than the emotional task at posterior sites.

3.2.3. *LPP*

The MANOVA for mean amplitude revealed a Hemisphere \times Longitudinality \times Task interaction ($F(2,144) = 6.23, p < 0.01, \eta_p^2 = .08$). Following up the interaction with separate ANOVAs for each level of Longitudinality revealed a significant Hemisphere \times Task interaction at posterior sites ($F(1,72) = 8.28, p < 0.01, \eta_p^2 = .10$) with a clear positivity for the emotional task that was larger for the left than the right hemisphere while for the linguistic task there was a much smaller positivity that was larger for the right than the left hemisphere. For frontal and central sites there was no significant Task \times Hemisphere interaction ($p > 0.05$).

Further, there was a Hemisphere \times Laterality \times Task interaction ($F(2,144) = 7.65, p < 0.01, \eta_p^2 = .10$) Follow-up ANOVAs for each task revealed that for the emotional task there was a significant Hemisphere \times Laterality interaction ($F(1,40) = 13.03, p < 0.01, \eta_p^2 = .25$) with a positivity for proximal sites that was larger for the right than left hemisphere and a negativity for lateral sites that was larger for the right than the left hemisphere, while for the linguistic task there was no significant interaction ($p > 0.05$).

To summarize, for the emotional task, there was a late positivity that was larger in the left than the right hemisphere at posterior sites and larger in the right than left hemisphere for proximal sites irrespective of intra-hemispheric location, while for the linguistic task there was a highly reduced positivity that was larger in the right than left hemisphere (see Figure 2).

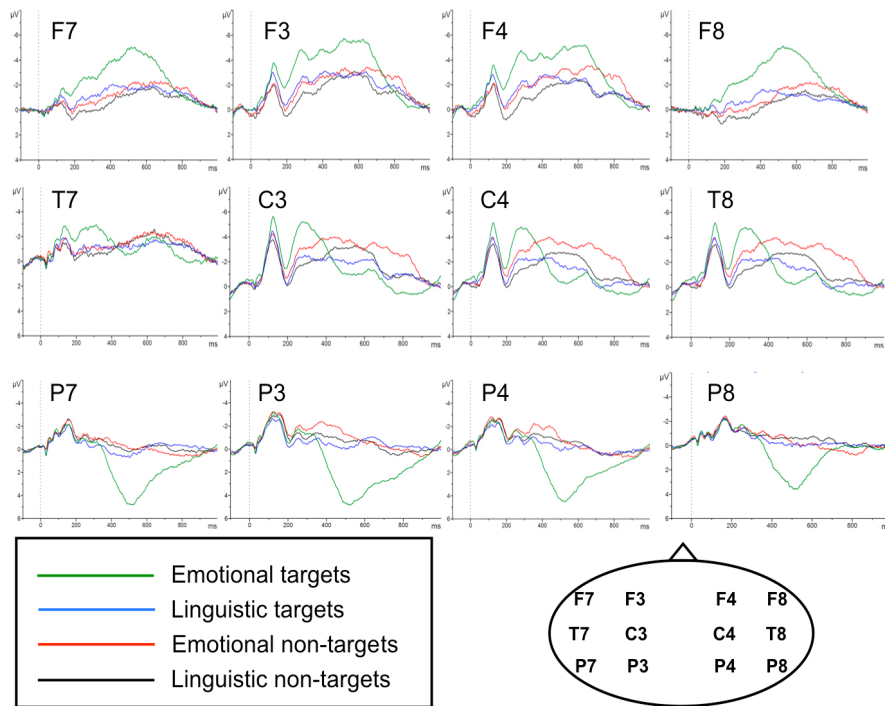


Figure 2. Grand average for the perception of emotional (targets: green, non-targets: red) and linguistic (targets: blue, non-targets: black) prosody for a representative set of electrodes used in the analyses. Averages are shown for a 1000 ms time window post target onset.

3.3. Electrophysiological results - ERP correlates of the ear advantage

The LEA and REA group consisted of 46 and 28 participants respectively. The two groups did not differ in male-female ratio, age or task performance (for all: $p > 0.05$). The groups showed a very large difference in the mean ear advantage $T(1,71) = -10.17$, $p < 0.001$, Cohen's $d = 2.46$) with a very large mean LEA and REA of -34 ms (SD = 26; Cohen's $d = -1.30$) and 33 ms (SD = 30; Cohen's $d = 1.14$), respectively. For both groups, the ear advantage was very robust (paired t-test for both groups; $p < 0.001$). In the section below, only significant interactions with EA-group are discussed (for a complete overview of effects, please see the Supplementary information).

3.3.1. Early negativity

No significant main effects or interactions were found for the MANOVA with mean amplitude as dependent variable.

The MANOVA with peak latency as dependent variable revealed a Hemisphere \times Laterality \times EA-group interaction ($F(1,69) = 4.96$, $p < 0.05$, $\eta_p^2 = .10$). Follow-up MANOVAs for each group showed that for the LEA-group the two-way interaction was non-significant ($p > 0.05$) but for the REA group there was a significant interaction ($F(1,26) = 11.75$, $p < 0.01$, $\eta_p^2 = .31$) with shorter latency for the left hemisphere than the right, but only at lateral sites (see Figure 3).

3.3.2. N2

No main effects or interactions were found.

3.3.3. LPP

No main effects or interactions were found.

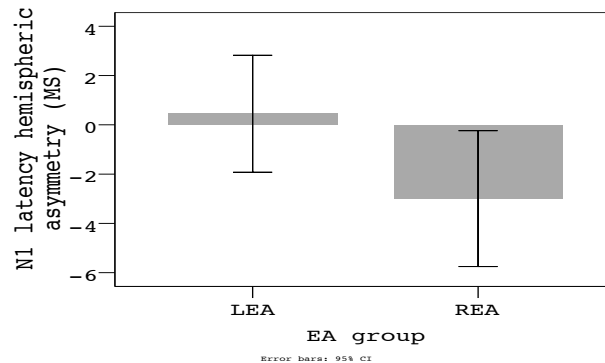


Figure 3. Hemispheric asymmetry of N1 peak latency (left hemisphere peak latency – right hemisphere peak latency) for lateral electrodes for the REA and LEA group.

4. Discussion

The goal of the present investigation was to test whether functional hemispheric specialization for prosody perception could be demonstrated. This was achieved by varying the emotional versus linguistic prosodic processing mode between participants using identical prosodic stimuli in a dichotic listening ERP paradigm. No functional lateralization effects were observed at the behavioral or electrophysiological level. Overall, the emotional task generated a larger response than the linguistic task. Last, the behavioral ear advantage correlated with hemispheric asymmetry of early negativity latency for the REA group. The absence of functional lateralization effects, task effects and the neural correlate of the right ear advantage will be discussed below.

4.1. Functional hemispheric specialization for prosody perception

As outlined in the introduction, recent meta-analytic evidence suggests that there is *relative* right hemispheric specialization for emotional prosody perception while linguistic prosody perception is under bilateral control (Witteman, Van IJzendoorn, Van de Velde, Van Heuven, & Schiller, 2011). Hemispheric specialization on the one hand has been proposed to result from a non-prosody specific right hemispheric advantage in the processing of relevant acoustic cues (i.e., the *cue-dependent* lateralization hypotheses) while the *functional* lateralization hypothesis on the other hand proposes that the left and right hemisphere are specialized in the processing of the *functional categories* of emotion versus linguistic prosody, respectively. The dynamic dual pathway model by Friederici and Alter (2004) further specifies that in the case of linguistic prosody perception, lateralization may depend on the degree to which prosody is bound to segmental structure, with linguistic prosody that is relatively closely bound to segments (such as metrical stress, as used in the present investigation) being left-lateralized and linguistic prosody that is not closely bound to the segmental structure (such as boundary marking) being right-lateralized. Note that the functional hypothesis requires that an abstract categorical level of processing has been reached (in order to deduce the functional category of prosodic information) before functional hemispheric specialization can take place.

As outlined in the introduction, recent models propose a three-stage process of (emotional) prosody perception (Kotz & Paulmann, 2011; Brück, Kreifelts, & Wildgruber, 2011), including (1) extraction of acoustic properties in the primary auditory cortex, (2) integration of acoustic properties into a meaningful suprasegmental representation in the associative auditory cortex, and (3) explicit evaluation in frontal cortical areas. It has been suggested that it takes at least 100 ms to reach the first more abstract stage (stage-two) level of processing (Schirmer & Kotz, 2006). Therefore, assuming a strictly serial model, functional lateralization effects are expected to be present in the electrophysiological signal 100 ms after the presentation of a prosodic stimulus. However, it seems likely that the prosody perception process is dynamic, allowing for the possibility of task demands modulating (hemispheric specialization for) earlier stages of prosodic perception (e.g., Brechmann & Scheich, 2005) through top-down modulation.

As both acoustic and functional properties have been hypothesized to influence hemispheric specialization for prosody perception (Pell, 1998) it is necessary to vary the functional task demands of a prosody perception task while keeping acoustics constant and to observe whether there is a shift in hemispheric asymmetry of neural activity to test the functional hemispheric lateralization hypothesis. Therefore, in

the present study we instructed one group of participants to categorize the emotional and another group to categorize the linguistic prosodic dimension of the same bidimensional prosodic stimuli. Furthermore, to maximize power to detect hemispheric specialization effects, we presented stimuli in a dichotic listening paradigm with a relatively large sample of participants as compared to previous research. Indeed, using G-power (Faul, Erdfelder, Lang, & Buchner, 2007) we calculated that even a small-to-moderate effect of the critical Ear \times Task interaction could be detected with sufficient (80%) power in the present study. However, no functional hemispheric specialization effects were found at the behavioral level (as indexed by the ear advantage) or at the electrophysiological level. Two explanations seem plausible for the absence of this effect: (1) functional lateralization effects are of considerable magnitude but the present paradigm was not able to detect the effect of interest; (2) functional lateralization effects are too small to be picked up even by the present (relatively high powered) study. The first explanation seems unlikely as a standard dichotic target detection paradigm was employed in the present investigation that has frequently been used in previous studies. Further, an ERP-waveform was detected that is typically observed in target detection paradigms (initial negativity followed by a positivity) demonstrating the validity of the paradigm and suggesting that the current paradigm should in principle have been able to detect functional lateralization effects.

Concerning the second explanation, an examination of the previous evidence in favor of functional lateralization is required. Only three previous studies had the necessary design to test the functional lateralization hypothesis. First, in a behavioral study, Luks, Nusbaum and Levy (1998) presented 32 subjects with utterances that were pronounced either as a question or a statement in the dichotic listening paradigm. Participants were instructed to categorize the utterances as a question or statement. No ear advantage was found. In a second experiment, 50 subjects had to categorize the same utterances but now emotionally (whether the utterances sounded surprised or neutral). This time a LEA was found. The authors concluded that the ear advantage could be modulated by task demands alone. However, these results are only partially in line with the functional lateralization hypothesis, as the expected REA for the linguistic prosody categorization task was not found. Second, Wildgruber et al. (2004) contrasted discrimination of sentential focus (linguistic prosody perception) to discrimination of expressiveness of the same stimuli (which can be argued to fall in the category of emotional prosody) in a functional imaging study including 10 participants. When contrasting the two tasks directly, a cluster of activity was observed in the left dorsolateral prefrontal cortex (DLPFC) for perception of linguistic prosody while bilateral orbitofrontal activation was noted for the perception of emotional prosody. Although this seems to provide evidence in favor of a greater left- than right-hemispheric contribution to linguistic prosody perception, no formal tests on hemispheric asymmetry were performed, leaving it unclear whether the left DLPFC was indeed significantly more active than its right-hemispheric counterpart. Also, the expected right lateralized activity for the emotional task was not found. Lastly, a recent ERP-study using the cross-splicing paradigm with a sample size of 20 participants also failed to find (functional) hemispheric specialization effects for emotional versus linguistic prosody perception (Paulmann, Jessen, & Kotz, 2012). Therefore, based on the scarce evidence to date, we conclude that the evidence for functional lateralization in prosody perception is weak and future high-powered studies that manipulate the

function of prosody independently of acoustics while measuring neural activity over the two hemispheres are required to clarify this issue.

4.2. *Task effects*

Although no significant interaction between task demands and hemispheric asymmetry was found, emotional task demands enhanced the amplitude of all ERP components across both hemispheres relative to the linguistic task. Two explanations may account for these effects. On the one hand, directing attention to the emotional dimension of the stimuli might have enhanced processing (or the emotional dimension might have interfered with linguistic processing during the linguistic task). Indeed, previous studies have found evidence for enhancement of ERP amplitudes by emotional task demands (e.g., for the LPP see Hajcak et al., 2006; Naumann et al., 1992), interpreted as prioritized processing of emotional information. On the other hand, task differences could be explained by differences in task difficulty, as the emotional task was easier than the linguistic task in the present investigation. Indeed, previous ERP studies have found enhanced amplitudes when difficulty of a task is reduced (e.g., for the LPP, see Molnár, 1999; Gaál, Csuha, & Molnár, 2007; but see Combs & Polich, 2006 for conflicting evidence). These two explanations might be non-mutually exclusive as processing of emotional information might be easier than linguistic processing because of its comparatively large salience. Future studies that vary emotional versus non-emotional task demands while controlling for difficulty (or vice versa) are needed to shed light on this issue.

Lastly, early negativity amplitudes were larger across both hemispheres for the left ear than the right for the emotional task, while the reverse was true for the linguistic task. Although it is tempting to interpret this interaction as an electrophysiological correlate of LEA and REA respectively, these effects were unrelated to behavioral EA effects.

4.3. *Electrophysiological correlates of the Ear Advantage (EA)*

Although we did not find an overall EA for linguistic or emotional prosody perception, the relatively large sample size of the present investigation permitted an additional analysis to explore the electrophysiological correlates of the EA. Participants could be divided into a REA or LEA group independently of whether the emotional or linguistic task was performed, allowing for an analysis of possible underlying electrophysiological determinants of the EA. This analysis revealed that only hemispheric asymmetry in the latency of the early negativity for lateral sites could account for REAs, with REAs associated with shorter latency in the left than right hemisphere. For LEAs, however, such a latency difference was not found.

The ‘structural model’ of the ear advantage proposed by Kimura (see Della Penna et al., 2007) suggests that EAs arise in the DL-situation because ipsilateral neural routes from the ear to the hemisphere are suppressed. When the specialized hemisphere is ipsilateral to the stimulated ear, the signal has to be transferred over the corpus callosum to reach it and reaction time is delayed. The present results are in line with the structural model as REAs (but not LEAs) were associated with a delayed response of the ipsilateral hemisphere. Our results suggest that REAs can be explained by relatively early specialization at the level of the auditory cortex of the contralateral hemisphere, as the locus of the auditory N1 has been suggested to lay in the auditory cortex (Sandmann et al., 2007). Such early ERP correlates of the EA have been

reported before in the literature (Eichele, Nordby, Rimol, & Hugdahl, 2005, but for conflicting evidence see Greenwald & Jerger, 2003). However, other studies have found later ERP components to be correlated with the EA (Ahoniska, Cantell, Tolvanen, & Lyytinen, 1993; Teder, Alho, Reinikainen, & Näätänen, 1993; Erhan, Borod, Tenke, & Bruder, 1998). As suggested by Erhan, Borod, Tenke and Bruder (1996), the N1 asymmetry in favor of the left hemisphere might not be related to prosodic processing but rather to the detection of phonemes. Thus, although speculative, the REAs found in the present study may represent relatively early left hemispheric specialization for phonetic processing during prosody processing.

4.4. Strengths and weaknesses

The present study is the first to systematically test the functional lateralization hypothesis of prosody perception using both behavioral evidence (EAs) and ERPs and with the highest statistical power to date. Despite relatively high statistical power, we failed to find evidence for functional lateralization, which we have argued may reflect the modest magnitude of this effect. The task differences found were confounded with a task difficulty effect, rendering it uncertain whether these differences reflect differences in the linguistic versus emotional mode of processing or differential difficulty of the tasks. Lastly, we included only negative emotions, restricting inferences about emotion effects to negative valence.

4.5. Conclusion

The present investigation did not find evidence for the functional lateralization hypothesis of prosody perception despite relatively high statistical power, suggesting that functional lateralization effects are small. Evidence was found in favor of the idea that REAs can be explained by a speed of processing advantage of the contralateral auditory cortex, in line with the structural model of dichotic listening.

ACKNOWLEDGEMENTS

We thank Jos Pacilly for doing programming work for this study.

5. References

- Ahonniska, J., Cantell, M., Tolvanen, A., & Lyytinen, H. (1993). Speech perception and brain laterality: the effect of ear advantage on auditory event-related potentials. *Brain and Language, 45*, 127-146.
- Becker, F., & Reinvang, I. (2007). Successful syllable detection in aphasia despite processing impairments as revealed by event-related potentials. *Behavioral and Brain Functions, 3*: 6.
- Brechmann, A., & Scheich, H. (2005). Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cerebral Cortex, 15*, 578-587.
- Briggs, K. E., & Martin, F. H. (2009). Affective picture processing and motivational relevance: Arousal and valence effects on ERPs in an oddball task. *International Journal of Psychophysiology, 72*, 299-306.
- Brück, C., Kreifelts, B., & Wildgruber, D. (2011). Emotional voices in context: A neurobiological model of multimodal affective information processing. *Physics of Life Reviews, 8*, 383-403.
- Combs, L. A., & Polich, J. (2006). P3a from auditory white noise stimuli. *Clinical Neurophysiology, 117*, 1106-1112.
- Cutler, A. (2005). Lexical stress. In D. B. Pisoni & R. E. Remez (Eds.). *The Handbook of Speech Perception* (264-289). Oxford: Blackwell.
- Davidson, R. J., & Hugdahl, E. K. (Eds.). (1995). *Brain Asymmetry*. Cambridge, MA: MIT Press.
- Della Penna, S., Brancucci, A., Babiloni, C., Franciotti, R., Pizzella, V., Rossi, D... Romani, G. L. (2007). Lateralization of dichotic speech stimuli is based on specific auditory pathway interactions: Neuromagnetic evidence. *Cerebral Cortex, 17*, 2303-2311.
- Eichele, T., Nordby, H., Rimol, L. M., & Hugdahl, K. (2005). Asymmetry of evoked potential latency to speech sounds predicts the ear advantage in dichotic listening. *Cognitive Brain Research, 24*, 405-412.
- 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.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.
- Fitzgerald, P. G., & Picton, T. W. (1983). Event-related potentials recorded during the discrimination of improbable stimuli. *Biological Psychology, 17*, 241-276.
- Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language, 89*, 267-276.
- Gaal, Z. A., Csuha, R., & Molnár, M. (2007). Age-dependent changes of auditory evoked potentials--effect of task difficulty. *Biological Psychology, 76*, 196-208.
- Gamble, M. L., & Luck, S. J. (2011). N2ac: An ERP component associated with the focusing of attention within an auditory scene. *Psychophysiology, 48*, 1057-1068.
- Greenwald, R. R., & Jerger, J. (2003). Neuroelectric correlates of hemispheric asymmetry: spectral discrimination and stimulus competition. *Journal of the American Academy of Audiology, 14*, 434-443.

- Grimshaw, G. M., Kwasny, K. M., Covell, E., & Johnson, R. A. (2003). The dynamic nature of language lateralization: effects of lexical and prosodic factors. *Neuropsychologia*, *41*, 1008-1019.
- Hajcak, G., Moser, J. S., & Simons, R. F. (2006). Attending to affect: Appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion*, *6*, 517-522.
- Kotz, S. A., & Paulmann, S. (2011). Emotion, Language, and the Brain. *Language and Linguistics Compass*, *5*, 108-125.
- Luks, T. L., Nusbaum, H. C., & Levy, J. (1998). Hemispheric involvement in the perception of syntactic prosody is dynamically dependent on task demands. *Brain and Language*, *65*, 313-332.
- Molnár, M. (1999). The dimensional complexity of the P3 event-related potential: Area-specific and task-dependent features. *Clinical Neurophysiology*, *110*, 31-38.
- Nager, W., Rosenthal, O., Bohrer, I., Teder-Sälejärvi, W. A., & Münte, T. F. (2001). Human event-related potentials and distraction during selective listening. *Neuroscience Letters*, *297*, 1-4.
- Naumann, E., Bartussek, D., Diedrich, O., & Laufer, M. (1992). Assessing cognitive and affective information processing functions of the brain by means of the late positive complex of the event-related potential. *Journal of Psychophysiology*, *6*, 285-298.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97-113
- Paulmann, S., Jessen, S., & Kotz, S. (2012). It's special the way you say it: An ERP investigation on the temporal dynamics of two types of prosody. *Neuropsychologia*, *50*, 1609-1620.
- Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. *Neuropsychologia*, *36*, 701-715.
- Polich, J., & Bondurant, T. (1996). P300 Sequence effects, probability, and interstimulus interval. *Physiology and Behavior*, *61*, 843-849.
- Sandmann, P., Eichele, T., Specht, K., Jäncke, L., Rimol, L. M., Nordby, H., & Hugdahl, K. (2007). Hemispheric asymmetries in the processing of temporal acoustic cues in consonant-vowel syllables. *Restorative Neurology and Neuroscience*, *25*, 227-240.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*, 227-256.
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*, 24-30.
- Sussman, E., Winkler, I., Huotilainen, M., Ritter, W., & Näätänen, R. (2002). Top-down effects can modify the initially stimulus-driven auditory organization. *Cognitive Brain Research*, *13*, 393-405.
- Teder, W., Alho, K., Reinikainen, K., & Näätänen, R. (1993). Interstimulus interval and the selective-attention effect on auditory ERPs: "N1 enhancement" versus processing negativity. *Psychophysiology*, *30*, 71-81.
- Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Papers in linguistics: international journal of human communication*, *13*, 201-277.

- Van Lancker, D., & Sidtis, 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.
- Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., & Ackermann, H. (2004). Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. *Cerebral Cortex, 14*, 1384-1389.
- Witteman, J., Van IJzendoorn, M. H., Van de Velde, D., Van Heuven, V. J., & Schiller, N. O. (2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: A meta-analysis of the lesion literature. *Neuropsychologia, 49*, 3722-3738.
- Witteman, J., Van Heuven, V. J., & Schiller, N. O. (2012). Hearing feelings: A quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia, 50*, 2752-2763.

Appendix A

Pseudowords used in the experiments:

dundon
kaldun
duldin
kanpal
daldan
kuldul
kuldil
duldun
dalpal
kulpul