

## **The Neurocognition of Prosody**

Silke Paulmann

Department of Psychology

Centre for Brain Science

University of Essex

6739 words (excl. references)

Correspondence:

Department of Psychology

University of Essex

Wivenhoe Park

Colchester

CO4 3SQ

UK

[paulmann@essex.ac.uk](mailto:paulmann@essex.ac.uk)

### Abstract

Prosody is one of the most undervalued components of language, despite fulfilling manifold purposes: it can, for instance, help assign the correct meaning to compounds such as ‘white house’ (linguistic function), or help a listener understand how a speaker feels (emotional function). However, brain based models that take into account the role prosody plays in dynamic speech comprehension are still rare. This is probably due to the fact that it has proven difficult to fully denote the neurocognitive architecture underlying prosody. This review discusses clinical and neuroscientific evidence regarding both linguistic and emotional prosody. It will become obvious that prosody processing is a multi-stage operation and that its temporally and functionally distinct processing steps are anchored in a functionally differentiated brain network.

## Introduction

Verbal communication is often a question of “tone”. Modulating parameters such as vocal pitch (high/low), loudness (loud/silent), tempo (fast/slow), or voice quality (clear/harsh), allow us to give the correct meaning to what we are saying and help the listener to interpret a message correctly. This is not only true when expressing how we feel (e.g. angry, nervous, happy), but also when conveying non-emotional information. For instance, by raising (or not raising) our voice when articulating a string of words we can alter the interpretation of an utterance (e.g. changing from a statement such as *You finished writing the chapter.* to a question *You finished writing the chapter?*). Suprasegmental parameters of speech (prosody) can also be used to convey lexical meanings (e.g. *hot dog* vs. *hot dog*), or discourse information (e.g. new information is often accented while old information is de-accented; prosodic phrasing guides syntactic sentence interpretation) next to expressing emotions and attitudes (e.g. often, a raised voice is associated with an angry speaker, whereas a lowered voice might indicate that the speaker feels sad). Thus, prosody serves several linguistic and non-linguistic (emotional) functions; however, it is often an undervalued component of spoken language and brain based models that take into account the role prosody plays in dynamic speech comprehension are still rare (but see Friederici & Alter, 2004) and controversially discussed. In fact, much of the controversy around the neural basis of prosody is probably due to the fact that it fulfils several communicative functions (often at once): thus, the question that has been driving past research is whether the different prosodic functions are in- or interdependent. This chapter reviews past research on each function and outlines our current understanding of the neurocognitive architecture underlying prosody.

### Brain Mapping of Prosody

Historically, investigations into the cerebral representation of prosody aimed to specify whether one hemisphere dominates control over linguistic and/or emotional prosody processing by looking at lesion data. These early *simple hemispheric models* (Sidtis and Van Lancker Sidtis, 2003) resulted in three main hypotheses:

- i. Prosody, irrespective of communicative function (e.g. linguistic, emotional), is lateralized to the right cerebral hemisphere. Interactions with other linguistic information such as syntactic or semantic information are mediated through the corpus callosum (e.g., Klouda, Robin, Graff-Radford, & Cooper, 1988; Friederici and Alter, 2004).
- ii. Emotional prosody is processed in the right hemisphere (e.g., Blonder, Bowers, & Heilman, 1991; Ross, 1981).
- iii. Linguistic prosody is processed in the left hemisphere while emotional or affective prosody processing can be linked predominantly to the right hemisphere (known as Functional Lateralization Hypothesis; Van Lancker, 1980).

Although it has since been shown that a simple cortical hemispheric distinction is not substantiated by the available data (for reviews see e.g. Baum & Pell, 1999; Sidtis and Van Lancker Sidtis, 2003), the effect of these proposals can still be felt when scanning through the more recent literature. Many researchers base their hypotheses about the brain network underlying prosody processing on the premise that linguistic and emotional prosody are fully distinct entities in the language system and that these processes are (each) fully lateralized to one cerebral hemisphere (but see Ameby Seddoh, 2002, arguing why emotional and linguistic prosody should not be considered distinct categories). However, an alternative to simple hemispheric models was put forward by Van Lancker & Sidtis (1992):

- iv. Hemispheric lateralization is based on physical features: pitch is preferably processed by right hemispheric brain structures, whereas duration and intensity are primarily processed by left hemispheric structures (known as the Cue or Physical Feature Dependent Hypothesis).

An additional problem with the early simple hemispheric models is that they neglect the role that subcortical brain regions play during prosody processing. However, accumulating evidence from lesion patients showed that impaired prosody processing is often associated with damage to structures such as the caudate nucleus, putamen, and/or globus pallidus (i.e. the basal ganglia (BG); e.g., Cancelliere and Kertesz, 1990; Breitenstein, Daum, & Ackermann, 1998; Paulmann, Pell, & Kotz, 2008; Paulmann, Pell, & Kotz, 2009). Hence, a fifth hypothesis posits that

- v. Prosody processing is heavily mediated by subcortical brain regions without a strong hemispheric lateralization.

In an attempt to consolidate the different hypotheses, Sidtis and Van Lancker Sidtis (2003) proposed a Neurobehavioral Approach to Dysprosody. This framework suggests that prosody processing is not driven by one single mechanism, but instead relies on a complex conglomerate of motor, perceptual, and more cognitively based functions. Thus, a widespread, bilateral brain network might be implicated in prosody processing and discrepancies across lesion studies are probably due to the fact that dysprosody can materialize after disruption to any of the involved mechanisms linked to different brain regions.

More recently, similar working models, that is frameworks that suggest a highly differentiated brain network underlying *emotional* prosody processing, have been put forward by researchers who based their hypotheses primarily on evidence obtained from neuro-imaging

studies instead of data from the lesion approach (e.g. Schirmer & Kotz, 2006; Kotz & Paulmann, 2011; Wildgruber, Ethofer, Grandjean & Kreifelts, 2009; Brück, Kreifelts, & Wildgruber, 2011). Moreover, some of these models also hypothesized the temporal dynamics of emotional prosody processing. For instance, Kotz & Paulmann (2011) suggest that an initial extraction of acoustic cues (e.g. fundamental frequency, loudness, voice quality) takes place within 100 ms of stimulus onset. This process is argued to be mediated by primary and secondary auditory cortices (bilaterally). Once acoustic properties have been accessed, derivation of emotional salience/meaning (established through integration of emotionally relevant acoustic cues) occurs within 200 ms after stimulus onset. This process has been linked to the right anterior superior temporal sulcus/superior temporal gyrus. Finally, more elaborate processes (e.g. integration of information from prosody with semantics or broad context) could start around 400 ms after stimulus onset. These higher cognitive processes are presumed to be mediated by inferior frontal and orbito-frontal cortex (bilaterally). Thus, researchers moved away from the assumption that prosodic processing is one single mechanism and instead suggest that different emotional prosody processing stages are subserved by different brain areas. Ideally, future models on the neural circuitry regulating emotional and linguistic prosody will be able to integrate findings from both lesion and imaging fields of the literature.

The following review aims to show how results from clinical and empirical neuroscience studies have helped shape our understanding of prosody processing. To provide an integrative view of key findings, evidence from important past and more recent studies will be discussed, followed by a summary of how available evidence supports prosody processing models.

## The Neural Basis of Linguistic Prosody Processing

### *Clinical Evidence*

Initially, scientists predominantly relied on the lesion approach to specify which brain regions might underlie prosody processing. For instance, one of the earlier studies to explore the contribution of the right hemisphere (RH) to linguistic prosody processing was conducted by Weintraub and colleagues (1981). They tested RH patients and healthy controls (HCs) and assessed their ability to discriminate between phonemic (e.g. DARK room vs. dark ROOM) and sentential stress (e.g. STEVE drives the car vs. Steve drives the CAR). Patients were found to be outperformed by controls on these tasks, prompting the authors to suggest that the RH plays a strong role in linguistic prosody perception. Around the same time, Baum and colleagues (1982) investigated phonemic and sentential stress comprehension in left hemisphere (LH) patients and HCs. Their results showed that LH patients *also* suffer from difficulties on these tasks, challenging the view that the RH alone is involved in linguistic prosody perception. Heilman et al. (1984) thus tested performance of *both* LH and RH patients on linguistic prosody processing in one study. The authors report difficulties in identifying the modality (question, statement, command) of filtered sentences by both LH and RH patients when compared to HCs. Again, these results challenged the view that the RH is solely responsible for linguistic prosodic processing and instead point to a possible additional involvement of the LH during linguistic prosody perception. In fact, evidence from a subsequent study comparing LH and RH patient performance on phonemic stress identification suggests that LH patients can even be more strongly impaired than RH patients (Emmorey, 1987). However, only two years later, conflicting evidence emerged when Bryan (1989) reported results from a range of tasks (e.g. sentential and phonemic stress identification, discrimination and identification of sentence modalities) which

showed that RH patients were outperformed by LH patients on most of the tasks administered. Although generally better than RH patients, LH patients performed poorly on the majority (but not all) of the tasks when compared to HCs suggesting that indeed both RH and LH patients show impairments for linguistic prosody processing. The importance of task effects was further looked at years later. Pell and Baum (1997) asked RH and LH patients as well as HCs to either identify or to discriminate between different sentence modalities (interrogative, declarative, imperative). Both patient groups performed comparable to the HC group when discriminating between different prosodic patterns. However, when looking at the identification task, patients performed significantly worse than HCs. Not only do these data point to the possibility that linguistic prosody is processed bilaterally in the brain, but they also nicely show that severity of impairment might depend on task instructions.

In the years that followed, researchers continued to get conflicting results when exploring the influence of the LH and RH on linguistic prosody processing: Borod and colleagues (1992) report findings obtained from LH and RH patients and HC. Non-sense syllable strings (e.g. pa-da-ka) which were intoned in three different ways (declarative, interrogative, emphatic) had to be discriminated by participants using multiple-choice response cards. Results showed that RH patients made significantly more errors than LH patients and HCs who did not differ from each other. In contrast, Breitenstein, Daum, and Ackermann (1998) report data from LH and RH patients who performed comparably to HC when discriminating between sentence modality pairs. Yet again, Pell (1998) showed that LH patients are worse at identifying emphatic stress patterns than RH patients or controls. Similarly, Walker, Daigle, and Buzzard (2002) showed that LH patients performed significantly worse than RH patients and HC when identifying lexical or sentential stress. In addition, LH patients also suffered from difficulties in prosodic phrasing. As



outlined above, task differences (e.g. instructions, complexity) between studies are likely to affect results; however, discrepancies between findings from clinical research can also likely be linked to differences in patients' lesion locations and size as well as to differences in their speech and language abilities (see e.g. Baum and Pell, 1999, or Kotz, Meyer, Paulmann, 2006, for similar observations). Thus, Seddoh (2006) subdivided his LH patient population into three subgroups: Wernicke's Aphasia patients, Broca's Aphasia patients, and Global Aphasia patients. Their ability to identify the sentence modality (question vs. statement) was assessed. Results confirmed that Wernicke's Aphasia patients had no difficulties with the task, while Broca's Aphasia patients suffered from difficulties in identifying questions and Global Aphasics suffered from difficulties in identifying statements and questions. Results suggest once more that the LH can be critically tied to linguistic prosody processing but also support the view that patients' lesion locations need to be controlled for better than by LH/RH distinctions. Latter conclusion is also underpinned by recent data by Rymarczyk and Grabowska (2007) who subdivided RH patients into three groups (patients with lesions to frontal, temporo-parietal, or subcortical brain structures). They looked at the performance of identifying as well as discriminating between sentence modalities and discriminating between empathic stress patterns. Although all patient groups perform significantly worse than HC on all three tasks, results confirm the importance of controlling for lesion location as patients with lesions to temporo-parietal sites performed worse than the two other groups. The view that we need to look at patients whose lesion delineation is comparable was followed by Kho and colleagues (2008). They looked at RH and LH patients who underwent anterior temporal cortex resection. The authors report no differences between RH and LH patients when detecting word or contrastive stress, or during sentential discrimination.

Taken together, the evidence elaborated above clearly shows that clinical evidence provides very little convergent evidence that linguistic prosody is processed solely by one hemisphere. Moreover, the evidence from lesion patients which suggests that not only cortical, but also subcortical brain structures play a critical role during prosody processing (Brådvik et al., 1991; Cancelliere & Kertesz, 1990; Ross & Mesulam, 1979; Starkstein et al., 1994; Rymarczyk and Grabowska, 2007) also challenges all simple hemispheric models, though the idea that prosody is primarily mediated by subcortical structures is also clearly not substantiated. In short, although telling, clinical evidence alone has not helped to provide support for either of the original hypotheses put forward. This conclusion is supported by a recent ALE meta-analysis by Witteman et al. (2011) who report that lesions to the LH or RH have similar detrimental effects on linguistic prosody perception.

### *Brain Imaging Evidence*

Over the past two decades, fMRI and PET techniques have become popular tools for observing normal brain function. Motivated by the heterogeneous results from clinical studies, several investigations have been conducted to further delineate the brain network underlying prosody processing. However, in contrast to emotional prosody, imaging studies on linguistic prosody processing have been rare. It is likely that the limited number of imaging studies on linguistic prosody are linked to the problem that linguistic prosody is an umbrella term that refers to studies exploring a variety of processes linked to supra-segmental changes including (but not limited to) sentence type, phrase boundary, word stress, and pitch contour processing. Some of the existing research will be summarized below to show that despite using more fine-grained

methodologies, the brain network underlying linguistic prosodic processing is still not fully specified.

In an early PET study, Gandour and colleagues (2000) investigated pitch perception of one syllable Thai words in speakers of tone and non-tone languages (Thai, Chinese, English). Their results suggest that neural mechanisms underlying pitch perception differ depending on linguistic relevance of stimuli. Specifically, left frontal operculum activity was found for Thai speakers when discriminating between pitch patterns but the same activity was not found for English speakers. Also, Thai speakers failed to show the same left lateralized activation when discriminating pitch patterns of non-speech stimuli, suggesting that “linguistic relevance” might modulate lateralized activation patterns. This hemispheric laterality effect was confirmed in an fMRI study a few years later (Gandour et al., 2004) when brain activity of Chinese and English speakers was measured. Participants had to discriminate between one and three syllable long utterances and four different Chinese tones. Chinese (i.e. the tone language group) but not English participants showed left lateralized activation in inferior parietal and posterior superior temporal, anterior temporal, and frontopolar brain regions. Both English and Chinese participants showed right lateralized activation hot spots in parts of the superior temporal sulcus as well as the middle frontal gyrus. Similar to the previous study (Gandour et al., 2000), the authors interpret their results to suggest that linguistic relevance/knowledge can influence brain activation patterns for prosody. In particular, the tone language group is argued to have an implicit understanding about the relationship between acoustic cues and internal representations of suprasegmental sentence information. In contrast, listeners from non-tone language backgrounds can not possess the same higher-order prosodic representations. Thus, it was argued

that LH lateralization is linked to higher-order prosodic processing while RH lateralization might represent lower-order acoustic feature processing (Gandour et al., 2004).

Imaging studies looking at sentences rather than syllables or words often report RH dominance for linguistic prosodic processing. For instance, Meyer et al. (2002) investigated the neuroanatomical correlates of slow prosodic modulations. They presented participants with delexicalized (i.e. stimuli that contain no lexical-semantic information) and normal speech. When comparing stimuli from both conditions, the authors found increased activation in the right superior temporal brain region as well as in the fronto-opercular cortex for filtered speech, suggesting a strong RH involvement during linguistic prosodic speech processing (and slow pitch movements in particular). Similar RH dominance in response to prosodic stimuli was reported by Plante, Creusere, and Sabin (2002) who also explored the neural correlates of sentential prosody. Participants were again presented with low-pass filtered (i.e. delexicalized speech) and normal speech and had to carry out tasks high (remember and recognize words) or low (no task) in memory load. When comparing hemodynamic responses for prosodic speech with responses to unfiltered speech in the “no task” condition, the authors found a stronger bilateral activation within the superior temporal gyrus for prosodic speech. Moreover, when looking at frontal lobe activation patterns, results showed that the tasks high in memory load affected processing of filtered and unfiltered speech differently: while both speech stimuli resulted in bilateral activation of the frontal lobes, processing filtered stimuli resulted in stronger RH activation than processing unfiltered stimuli.

Influence of task instructions on prosodic processes was also reported by other authors who looked at sentence material (e.g. Wildgruber et al., 2004; Geiser, Zaehle, Jancke, & Meyer, 2008) albeit they failed to confirm a strongly right lateralized network for linguistic prosodic

processes. Instead, they showed that participants who focus on linguistic prosodic aspects of stimuli exhibit greater activation in the LH. Specifically, in their fMRI study, Wildgruber and colleagues (2004) showed right lateralized activation of the dorso-lateral frontal cortex and bilateral activation of thalamic and temporal regions in contrast to a rest condition for stimuli which were synthetically manipulated to exhibit different sentence foci (second word vs. final word focus). However, when participants were asked to focus on linguistic characteristics of stimuli, activation of the left inferior frontal gyrus was reported (as opposed to activation of bilateral orbito-frontal areas when the task focus was emotional). Similarly, Tracy et al. (2011) report that processing pitch information from lexical stimuli is predominantly recruiting left lateralized structures including the cingulate gyrus, middle temporal and superior temporal gyri, whereas pitch processing from tone sequences relied on right frontal and temporal cortices.

Taken together, data from neuro-imaging studies clearly suggest that a complex neural network spanning both hemispheres underlies linguistic prosodic processing. Findings further imply that various factors can impact on lateralization of effects. Specifically, it has been shown that lateralization of prosody can depend on task focus (linguistic/non-linguistic, active/passive; e.g. Wildgruber et al., 2004), task demands (high/low; e.g. Plante et al., 2002), language background/experience (tone/non-tone language; e.g. Gandour et al., 2000), acoustic cue (pitch/duration; e.g. Van Lancker & Sidtis, 1992), stimulus type (syllable/word/sentence; e.g. Meyer et al., 2002; Gandour et al., 2004), as well as methodological factors including design (event-related/blocked; c.f. Kotz, Meyer, Paulmann, 2006) and contrast/comparison conditions (rest/alternative prosodic function) .

*ERP findings*

Next to exploring the brain structures involved, research has also tried to specify the time-course of linguistic prosody using event-related brain potentials (ERPs). The temporal online dynamics of prosody are of particular interest as information might consolidate the different proposals on how prosody is represented in the brain considering that research suggests that *emotional* prosodic processing is a multi-stage process (e.g. Schirmer and Kotz 2006; Kotz & Paulmann, 2011; Wildgruber et al. 2006, 2009) with each stage linked to different brain areas. It is likely that a similar multi-layered mechanism applies to linguistic prosody processing.

Several ERP components with different onset latencies have been described. For instance, Steinhauer et al. (1999) explored how prosody helps listeners to establish a syntactic structure during language comprehension. They were the first to report that the so-called closure positive shift (CPS) is elicited quickly after prosodic phrase boundaries (which usually coincide with syntactic boundaries). This marker of prosodic boundary processing is also found in delexicalized speech (e.g. Steinhauer & Friederici, 2001), as well as during implicit prosodic processing situations (e.g. reading; Hwang & Steinhauer, 2011). Generally speaking, the CPS is elicited rapidly (between -100 and 0 ms) after the offset of a pre-boundary word and is thus argued to be triggered by preboundary syllable lengthening (e.g. Pauker et al., 2011). A slightly later prosody related component was recently reported by Li and colleagues (2011) who observed a fronto-centrally distributed negative ERP between 270-510 ms in response to prosodic prominence manipulations. Responses to prosodic boundary violations resulted in a longer negative ERP effect lasting from 270 ms to 660 ms. For a similar time-window, Böcker, Bastiaansen, Vroomen, Brunia, and de Gelder (1999) found that extraction of metrical stress from bi-syllabic words can be linked to the N325 which is elicited under both active (discrimination) and passive (listening) tasks.

In addition, processing of prosodic contour expectancy violations has been studied. For instance, Paulmann et al. (2012) recently violated linguistic prosodic expectancy by merging the beginning of a declarative sentence with the end of a question. Results revealed a frontally distributed Prosodic Expectancy Positivity (PEP) 620 ms after the onset of prosodic violations irrespective of task focus (linguistic/emotional), suggesting that listeners can detect an abrupt change in prosodic contour and that discrepancies in contours are quickly reanalyzed (Paulmann, Jessen, Kotz, 2012). Similar late positivities were found by Astésano et al., (2004) who report a P800 for similar prosodic expectancy violations and Eckstein & Friederici (2005) who report a P600 response to prosodic incongruity of the final word of a sentence.

These findings lend support to the assumption that participants not only use prosody to build up information about the sentence structure and modality but also realize quickly if the expectation is not fulfilled. However, these prosodic processes can sometimes be influenced by task focus as the P800 in Astésano et al.'s study was only observed when participants focused on the prosody, while the positivity reported by Eckstein & Friederici (2005) was found even without participants' explicit focus on prosody. In sum, ERP results confirm that prosody interfaces with other language functions such as semantics (e.g. Paulmann et al., 2012) and syntax (e.g. Eckstein & Friederici, 2005; Steinhauer et al., 1999) during online language comprehension. Given the different distributions of prosody related ERP effects and their differing temporal dynamics, it seems reasonable to assume that linguistic prosody processing is not only multi-faceted but at least partly hinges on differing neural mechanisms depending on the precise function which could recruit brain structures from both hemispheres.

## **The neural basis of emotional prosody processing**

### *Clinical Evidence*

Emotional prosody has long been of special interest to researchers exploring hemispheric specialization of brain processes, probably because this type of prosodic function cannot easily be associated with pure emotion (historically linked to the RH, e.g. Borod et al., 1998a,b) or pure language (historically linked to the LH, e.g. Friederici, 2002) processing. Instead, emotional prosody is at the intersection of both domains. Early research on how emotional prosody is anchored in the brain was done by Heilman, Scholes, and Watson (1975). They asked patients with right and left temporo-parietal lesions to identify the emotional tone a speaker used when intoning semantically neutral sentences. Results revealed that RH patients performed significantly worse on the task than LH patients suggesting that emotional prosody is primarily processed in the RH, though the lack of a HC group posits a problem to this conclusion. Thus, a few years later, the same group (Heilman et al., 1984) investigated the comprehension of emotional prosody in RH and LH patients and this time compared their performance to HCs. Results confirmed a deficit for RH patients when compared to LH patients and HCs. Similarly, Bowers, Coslett, Bauer, Speedie, and Heilman (1987) report that RH patients performed significantly worse than LH patients and HC when discriminating between emotional categories of prosodically and semantically emotional congruent and incongruent sentences as well as low-pass filtered sentences. This once more indicated an important role for the RH in emotional prosody perception. Results from Blonder et al. (1991) also revealed that RH patients have difficulties discriminating amongst different emotional categories for emotionally intoned neutral sentences when compared to LH and HC. Although subsequent evidence has often confirmed the RH involvement in emotional prosody perception (e.g. Blonder et al., 1991; Lalande et al., 1992;



Borod et al. 1998a,b; Ross & Monnot, 2008; Rymarczyk & Grabowska, 2007), there is some clinical data that questions the unique role of the RH (e.g. Schlanger et al., 1976; Breitenstein et al., 1998; Starkstein et al., 1994; Cancelliere & Kertesz, 1990).

For instance, Schlanger, Schlanger, and Gerstman (1976) described no differences between LH and RH patients for identifying the emotional tone of a speaker, suggesting that LH brain structures can also play a role during emotional prosody perception. Van Lancker and Sidtis (1994) asked LH and RH patients to identify emotional prosodic speech samples by matching them to facial expressions and emotional labels. They also fail to report differences between groups with respect to emotional prosody perception, indicating that emotional prosody perception is mediated through a bilateral network of brain structures. Support for this hypothesis comes from recent work by Kho and others (2008) who showed that both RH and LH temporal lobe epilepsy patients were impaired on emotional prosody recognition when compared to HC. Similarly, Starkstein and colleagues (1994) report that both LH and RH patients displayed emotional prosody comprehension difficulties; though it should be noted that more detailed analyses also showed that RH patients with lesions in the BG and tempo-parietal cortex were most severely affected. Corroborating findings that subcortical brain structures are implied in an emotional prosodic network, Cancelliere and Kertesz (1990) report emotional prosody processing difficulties in patients with LH and RH lesions involving the BG. Paulmann and colleagues (2008) also report emotional prosody recognition deficits in patients with LH lesions in the BG (for a detailed review on the role of the BG in emotional prosody processing see Kotz, Hasting, Paulmann, 2013), rendering it unlikely that emotional prosody perception is uniquely mediated through RH cortical regions. In fact, in an attempt to illuminate how the LH and RH might contribute differently to emotional prosody recognition, Van Lancker and Sidtis (1992)

explored whether patients with lesions in either hemisphere made similar errors in an emotional prosody identification task. Their findings suggest that patients used acoustic cues differently when judging the emotional tone of the speaker (Van Lancker & Sidtis, 1992). Specifically, a discrimination analysis on error patterns revealed that patients with lesions in the LH relied on pitch information to infer emotionality of stimuli, while patients with lesions in the RH seemed to predominantly use durational cues to infer emotionality. These data were argued to support the physical feature dependent hypothesis; however, following the same methodology, Pell & Baum (1997) failed to find differences between LH and RH patients in their misclassifications of stimuli.

In short, results from clinical studies do not support simple hemispheric models but instead suggest that emotional prosody processing recruits a broad network of cortical and subcortical brain regions possibly slightly more right than left lateralized. The modestly bigger RH involvement is corroborated by a recent meta-analysis (Witteman et al., 2011). When comparing results from studies testing RH and LH patients directly, findings implied that damage to the RH results in more severe problems for emotional prosody perception than damage to the LH. Finally, task focus, stimuli differences, differences in experimental paradigms, and, most critically, differences in patients' lesion size and location can heavily impact on findings and thus might contribute to equivocal findings.

### *Brain Imaging Evidence*

In line with studies reviewed, neuroimaging findings also tend to reveal that emotional prosody perception is mediated by a complex bilateral network involving both cortical and subcortical brain structures with some studies suggesting a slightly more right lateralized network. For

instance, an early PET study explored emotional prosody categorization of sentences (George et al., 1996). Results revealed significant blood flow changes in the right pre-frontal cortex. Similarly, Buchanan et al. (2000) asked participants to listen to emotionally intoned words while either detecting the emotionality of the speaker (emotion task), or detecting a probe word (non-emotion task). They report right lateralized frontal lobe and right lateralized anterior auditory cortex activation for emotionally intoned words when comparing the emotion task with the non-emotion task; however, a bilateral activation is reported for emotionally intoned words when comparing the emotion task to a rest baseline. Also, Mitchell and colleagues (2003) report right-lateralized activation of superior and middle temporal gyri when participants listened to emotional sentences spoken in different emotions. A right-lateralized network for emotional prosody perception was also confirmed by Beaucousin et al. (2007) who report greater activation in the right temporal lobe for emotional prosody processing when comparing emotional speech with text-to-text speech lacking emotional attributes. Similarly, Wildgruber et al. (2005) outline that recognition of emotionally intoned and semantically neutral sentences resulted in right lateralized activation of the posterior superior temporal sulcus, as well as dorsolateral, and orbitobasal frontal areas when comparing activation patterns for an emotional recognition with a phonetic monitoring task. However, when comparing activation patterns for the emotional recognition task with a rest baseline, frontal, temporal, and parietal brain areas were activated bilaterally, demonstrating again that task effects (and/or condition comparisons) play a crucial role when looking at brain activation patterns for emotional prosody comprehension (also see paragraph below).

In fact, as to be expected based on neuropsychological findings, bilateral brain activation is also reported frequently. For instance, Kotz et al. (2003) compared sentences spoken in an emotional tone of voice with filtered sentences that contained only prosodic information. Comparisons revealed a bilateral frontal and subcortical (BG) activation pattern for the emotional prosodic condition. In addition, Sander

and colleagues (2005) compared angry and neutral prosody processing in a dichotic listening paradigm. Irrespective of whether participants attended to or ignored the meaningless angry speech stimuli, right amygdala and bilateral superior temporal sulcus activation was reported. However, the same stimuli elicited greater bilateral activation in the orbitofrontal cortex and the cuneus (in the medial occipital cortex) when participants paid attention to presented stimuli as opposed to when they ignored them. Moreover, greater bilateral activation of mid-superior temporal sulcus has been reported for angry in contrast to neutral prosody (Grandjean et al. 2005). Also, Wiethoff et al. (2008) investigated activation for emotionally arousing prosody during passive listening. Specifically, they looked at happy, erotic, fearful, and angry prosody, thus included stimuli of both positive and negative valence. The authors report greater activation for arousing prosody in the right primary auditory cortex, the mid-superior temporal gyrus and the left temporal pole as well as the hypothalamus. Leitman and others (2010) tried to compare stimuli which were either rich or low in emotional acoustic cue saliency, i.e. stimuli which are easily recognizable by a single acoustic parameter such as fundamental frequency or intensity. They report greater activation for emotional prosodic stimuli that are rich in cue saliency for the planum temporale, posterior superior temporal and middle gyri (i.e. superior temporal cortex), as well as the amygdala while participants engaged in an emotional sentence identification task. In contrast, greater activation was found in inferior and temporo-frontal areas when participants processed stimuli with less salient or dominant acoustic cues. Interestingly, their results indicate that lateralization can depend on the specific emotion investigated, as stronger left lateralized activation patterns were found for angry prosody in contrast to more right lateralized hot spots for fearful and happy prosody.

In addition to exploring the influence of individual acoustic cues (or their saliency) on activation patterns, more recent research has also tried to illuminate the role of task effects in neural responses linked to emotional prosody processing. For instance, Bach and others (2008) compared activation patterns for emotionally intoned pseudo words while participants either had to decide which gender the speaker voice was or which emotion the speaker was trying to convey. They found bilateral activation of

the amygdala, left superior temporal sulcus and right parietal areas when participants judged the gender of the speaker while activation for left inferior frontal gyrus, bilateral parietal, anterior cingulate and supplemental motor cortex was found when participants focused on the emotionality of the stimuli. Moreover, when comparing emotional with neutral prosody, neural activity was found in right superior temporal gyrus, left inferior frontal gyrus, as well as in the anterior cingulate (bilateral), the insula and the putamen (bilateral). Interestingly, subcortical brain activation was particularly strong when participants focused on the emotionality of the stimuli, once more leading to the impression that task effects can critically impact on activation patterns. Finally, Ethofer et al. (2009) presented words spoken in an angry or neutral prosody to participants while they engaged in a valence discrimination or a word classification task. Activation in response to angry prosody was not only found in voice sensitive areas of temporal cortices, but also in the amygdala, insula, and mediodorsal thalami irrespective of which task participants engaged in. However, when comparing the valence discrimination with the word classification task, Ethofer and colleagues (2009) found stronger activation in the right middle temporal gyrus as well as in the orbito-frontal cortex (bilateral), suggesting that evaluation of emotional aspects activates these areas in particular.

In sum, neuroimaging data generally confirm that emotional prosody processing involves a bilateral temporo-frontal brain network, with some studies describing activation of subcortical structures. Activation hot spots between studies seem to differ depending on task focus and stimuli quality on activation patterns. Latter acoustic cue influence is confirmed in a meta-analysis by Witteman and colleagues (2012) who advocate that higher activation likelihood of the RH might stem from lateralized activation of primary and secondary auditory cortices (i.e. sensory cue processing).

### *ERP findings*

The time-course underlying emotional prosody processing has been explored in recent electrophysiological studies. Several ERP components have been of special interest: The N100 is

generally assumed to reflect processing of frequency (e.g. pitch) and loudness information, i.e. it is linked to the extraction of acoustic cues. This early component is followed by the P200, a fronto-centrally distributed component peaking 200 ms after stimulus onset (i.e. far before a sentence is completed) and that has been shown to be responsive to emotional prosodic (e.g. Paulmann et al., 2008; Schirmer, Chen, Ching, Tan, & Hong (2013) and arousal (Paulmann & Kotz, 2013) attributes of stimuli. Specifically, it has been outlined that different emotional prosodies can be distinguished from one another (Paulmann & Kotz, 2013) and from neutral (e.g. Paulman & Kotz, 2008, Schirmer et al., 2013) within 200 ms of stimuli onsets. Although the P200's sensitivity to pitch (Pantev et al. 1996) and loudness (Picton et al. 1977) variations has been demonstrated, research on emotional prosody implies that listeners rely on more than just one acoustic parameter when detecting emotional salience from auditory stimuli. However, which specific acoustic cue (configurations) are needed to detect the valence or even one particular emotion from speech still awaits further clarification. While the P200 is elicited under attentive processing conditions (albeit irrespective of implicit or explicit emotional tasks), the mismatch negativity has been linked to emotional category change detection under pre-attentive processing conditions (e.g. Schirmer, Striano, Friederici, 2008). Both components have repeatedly been linked to early emotional salience detection (e.g. Schirmer & Kotz, 2006; Paulmann & Kotz, 2008; Kotz & Paulmann, 2011).

This early emotional evaluation or appraisal of vocal expressions is followed by more elaborated stimulus evaluations. In particular, meaning evaluation or access to (emotional) memory representations have been linked to later ERP components such as the P300 (e.g., Wambacq and Jerger, 2004), N300 (e.g., Bostanov and Kotchoubey, 2004), N400 (e.g., Schirmer et al., 2002, 2005; Schirmer and Kotz, 2003; Paulmann and Pell, 2010), and the late positive complex (LPC) (Schirmer et al., 2013; Paulmann & Kotz, 2013). For instance, Bostanov and Kotchoubey (2004) assessed how emotional meaning is extracted from exclamations such as 'Wow' and 'Oooh' by presenting participants with emotionally congruent or incongruent stimuli. They report an enhanced N300 in response to

incongruous exclamations suggesting that emotional prosodic meaning is extracted around 300 ms after stimulus onset. Similarly, larger N400 amplitudes were found for prosodically/semantically incongruent emotional words (e.g. 'happy' spoken in an angry voice; Schirmer and Kotz, 2003). To explore how *much* prosodic information is needed to infer emotional meaning, Paulmann and Pell (2010) presented participants with emotionally intoned sentence fragments which were either 200 ms or 400 ms long. Sentence fragments served as primes and were followed by emotionally matching or mismatching facial expressions. N400-like priming effects were found for faces that were preceded by emotionally mismatching sentence fragments, though priming from shorter fragments led to a reversed effect. The findings nicely showed that listeners can extract emotions from both short and somewhat longer sentence fragments, supporting the view that 200 ms are sufficient to build up emotional context (c.f. P200 results above) and that emotionally relevant cues are extracted rapidly. Moreover, the N400-like priming effects are in line with studies proposing that emotional meaning is processed around 300 - 400 ms after stimulus onset. Finally, two recent studies report differently modulated LPCs in response to sentences differing in emotional tones, suggesting continued exhaustive processing of emotional prosodic information at late processing stages (Paulmann et al., 2013; Schirmer et al., 2013). Moreover, Schirmer and colleagues (2013) further showed that P200 modulations can predict modulation of subsequently elicited LPCs, suggesting that successful early emotional salience detection goes hand in hand with later more in-depths processing of emotional prosody. Arguably, latter step is necessary to ensure appropriate social behaviour.

Finally, to investigate in how far specific brain areas can be linked to individual emotional prosody processing steps, ERP lesion studies have been carried out (Paulmann et al., 2008; 2009; 2010). In a nutshell, these studies revealed that early emotional prosodic appraisal as reflected in the P200 component does not seem to be critically tied to the BG or orbito-frontal cortex. In contrast, later meaning-related processes (as reflected in the PEP and N400-like components) seem to be affected by lesions to the BG (Paulmann et al., 2008; 2009) and later emotional prosody recognition processes (as reflected in behavioural responses such as emotion recognition rates) are affected by lesions to the BG (Paulmann et al. 2008; 2009) and the orbito-frontal cortex (Paulmann et al. 2010). These findings highlight the potential of ERP lesion studies as they allow exploring the function of specific brain areas with methodologies that have excellent temporal resolution. Thereby they help illuminating which processing stages might be modulated via specific neural structures and this methodology can thus help to consolidate conflicting findings from lesion and neuroimaging findings as the latter two lack the high temporal resolution of ERPs.

The studies reviewed in this section thus confirm that emotional prosodic processing includes processes such as rapid early emotional appraisal as well as comprehensive emotional meaning processing, that these processes occur irrespective of directed attention of listeners, and that different functions can be linked to different underlying brain areas. Moreover, comparing findings described in this section (emotional prosody) and the previous one (linguistic prosody), it also seems as if the two processes generally elicit different ERPs (with different onset latencies). Indeed, when exploring the comparative nature of the time-course linked to the two functions, results from a recent ERP study (Paulman et al., 2012)



revealed that emotional prosodic expectancy violations are detected ~150 ms earlier than expectancy violation of linguistic prosody and that both violations resulted in different PEP distributions: linguistic PEPs were elicited predominantly at anterior electrode sites while emotional PEPs were most dominant at posterior electrode sites, suggesting that at least partly different neural mechanisms are at play during emotional and linguistic prosody processing. Taken together, and similar to the imaging literature reviewed above, differences in ERP studies (e.g. ERP polarity and latency) seem to be influenced by stimuli (e.g. words/sentences, normal/filtered/pseudo-speech), tasks (e.g. implicit/explicit emotional evaluation), and designs (e.g. blocked/randomized). However, despite methodological differences across studies, electrophysiological research has helped delineate the time-course underlying emotional prosodic processing clearly supporting the idea that emotional prosody processing is made up of different sub-processes.

## **Summary**

Years of research have shown that initially advocated simple hemispheric models fail to adequately describe brain mechanisms underlying prosody processing. Research has revealed that neural mechanisms of prosody are vulnerable to external influences such as task demands, stimulus quality, and experimental design thereby explaining some of the discrepant literature reports. Given the functional complexity of linguistic prosody processing, its neural specifications seem to be less clearly delineated than neural structures for emotional prosody. However, clinical and imaging results seem to suggest that both functions of prosody cover different sub-processes which are each anchored in different parts of the brain. Future development of brain based language models requires that the field continues to move away from

understanding prosody as a holistic concept and focuses instead on portraying each function as multi-layered.

### **Acknowledgements**

I would like to thank Rick Hanley for helpful comments on an earlier version of this chapter.

## References

- Amebu Seddoh, S. (2006). Basis of intonation disturbance in aphasia: Perception. *Journal of Neurolinguistics*, 19(4), 270-290.
- Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, 18, 172–184.
- Bach, D. R., Grandjean, D., Sander, D., Herdener, M., Strik, W. K., & Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *Neuroimage*, 42(2), 919-927.
- Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, 13(8), 581-608.
- Baum, S. R., Daniloff, J. K., Daniloff, R., & Lewis, J. (1982). Sentence comprehension by Broca's aphasics: Effects of some suprasegmental variables. *Brain and Language*, 17(2), 261-271.
- Beaucousin, V., Lacheret, A., Turbelin, M. R., Morel, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2007). fMRI study of emotional speech comprehension. *Cerebral cortex*, 17(2), 339-352.
- Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The role of the right hemisphere in emotional communication. *Brain*, 114(3), 1115-1127.
- Böcker, K. B., Bastiaansen, M., Vroomen, J., Brunia, C. H., & Gelder, B. (1999). An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology*, 36(6), 706-720.
- Borod, J. C., Cicero, B. A., Obler, L. K., Welkowitz, J., Erhan, H. M., Santschi, C., Grunwald, I.S., Agosti, R.M., & Whalen, J. R. (1998). Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology*, 12(3), 446.
- Borod, J.C., R. L. Bloom, & Santschi-Haywood, C. (1998b). Verbal aspects of emotional communication. Right hemisphere language comprehension: perspectives from cognitive neuroscience, ed. by M. Beeman and C. Chiarello, 285–307. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: Continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology*, 41(2), 259-268.
- Bowers, D., Coslett, H., 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.
- Brådvik, B., Dravins, C., Holtås, S., Rosen, I., Ryding, E., & Ingvar, D. H. (1991). Disturbances of speech prosody following right hemisphere infarcts. *Acta Neurologica Scandinavica*, 84(2), 114-126.
- Breitenstein, C., Daum, I., & Ackermann, H. (1998). Emotional processing following cortical and subcortical brain damage: contribution of the fronto-striatal circuitry. *Behavioural neurology*, 11(1), 29-42.
- Brück, C., Kreifelts, B., & Wildgruber, D. (2011). Emotional voices in context: a neurobiological model of multimodal affective information processing. *Physics of Life Reviews*, 8(4), 383-403.

- Bryan, K. L. (1989). Language prosody and the right hemisphere. *Aphasiology*, 3(4), 285-299.
- Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jäncke, L. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research*, 9(3), 227-238.
- Cancelliere, A. E., & Kertesz, A. (1990). Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and cognition*, 13(2), 133-147.
- Eckstein, K., & Friederici, A. D. (2006). It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience*, 18(10), 1696-1711.
- Emmorey, K. D. (1987). The neurological substrates for prosodic aspects of speech. *Brain and Language*, 30(2), 305-320.
- Ethofer, T., Van De Ville, D., Scherer, K., & Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Current Biology*, 19(12), 1028-1033.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, 6(2), 78-84.
- Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: a dynamic dual pathway model. *Brain and language*, 89(2), 267-276.
- Gandour, J., Tong, Y., Wong, D., Talavage, T., Dziedzic, M., Xu, Y., ... & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage*, 23(1), 344-357.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of cognitive neuroscience*, 12(1), 207-222.
- Garrido-Vásquez, P., Pell, M. D., Paulmann, S., Strecker, K., Schwarz, J., & Kotz, S. A. (2012). An ERP study of vocal emotion processing in asymmetric Parkinson's disease. *Social Cognitive and Affective Neuroscience*.
- Geiser, E., Zaehle, T., Jancke, L., & Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *Journal of Cognitive Neuroscience*, 20(3), 541-552.
- George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K. M., Herscovitch, P., & Post, R. M. (1996). Understanding emotional prosody activates right hemisphere regions. *Archives of neurology*, 53(7), 665.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature neuroscience*, 8(2), 145-146.
- Heilman, K. M., Bowers, D., Speedie, L., & Coslett, H. B. (1984). Comprehension of affective and nonaffective prosody. *Neurology*, 34(7), 917-917.
- Heilman, K. M., Scholes, R., & Watson, R. T. (1975). Auditory affective agnosia. Disturbed comprehension of affective speech. *Journal of Neurology, Neurosurgery & Psychiatry*, 38(1), 69-72.

- Hwang, H., & Steinhauer, K. (2011). Phrase Length Matters: The Interplay between Implicit Prosody and Syntax in Korean “Garden Path” Sentences. *Journal of cognitive neuroscience*, 23(11), 3555-3575.
- Kho, K. H., Indefrey, P., Hagoort, P., Van Veelen, C. W. M., Van Rijen, P. C., & Ramsey, N. F. (2008). Unimpaired sentence comprehension after anterior temporal cortex resection. *Neuropsychologia*, 46(4), 1170-1178.
- Klouda, G. V., Robin, D. A., Graff-Radford, N. R., & Cooper, W. E. (1988). The role of callosal connections in speech prosody. *Brain and Language*, 35(1), 154-171.
- Kotz, S. A., & Paulmann, S. (2011). Emotion, language, and the brain. *Language and Linguistics Compass*, 5(3), 108-125.
- Kotz, S. A., Hasting, A. S., & Paulmann, S. (2013). On the orbito-striatal interface in (acoustic) emotional processing. *The Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man*, 229.
- Kotz, S. A., Meyer, M., & Paulmann, S. (2006). Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design. *Progress in brain research*, 156, 285-294.
- Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and language*, 86(3), 366-376.
- Lalande, S., Braun, C. M. J., Charlebois, N., & Whitaker, H. A. (1992). Effects of right and left hemisphere cerebrovascular lesions on discrimination of prosodic and semantic aspects of affect in sentences. *Brain and language*, 42(2), 165-186.
- Leitman, D. I., Wolf, D. H., Ragland, J. D., Laukka, P., Loughhead, J., Valdez, J. N., Javitt, D.C., Turetsky, B.I., & Gur, R. C. (2010). “It's Not What You Say, But How You Say It”: a reciprocal temporo-frontal network for affective prosody. *Frontiers in human neuroscience*, 4.
- Li, X., Chen, Y., & Yang, Y. (2011). Immediate integration of different types of prosodic information during on-line spoken language comprehension: An ERP study. *Brain Research*, 1386, 139–152.
- Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human brain mapping*, 17(2), 73-88.
- Mitchell, R. L., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, 41(10), 1410-1421.
- Pantev, C., T. Elbert, B. Ross, C. Eulitz, and E. Terhardt. 1996. Binaural fusion and the representation of virtual pitch in the human auditory cortex. *Hearing Research*, 100, 164–70.
- Pauker, E., Itzhak, I., Baum, S. R., & Steinhauer, K. (2011). Effects of cooperating and conflicting prosody in spoken English garden path sentences: ERP evidence for the Boundary Deletion Hypothesis. *Journal of cognitive neuroscience*, 23(10), 2731-2751.
- Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport*, 19(2), 209-213.

- Paulmann, S., & Pell, M. D. (2010). Contextual influences of emotional speech prosody on face processing: How much is enough?. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 230-242.
- Paulmann, S., Bleichner, M., & Kotz, S. A. Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology*, *4*, 345.
- Paulmann, S., Jessen, S., & Kotz, S. A. (2012). It's special the way you say it: An ERP investigation on the temporal dynamics of two types of prosody. *Neuropsychologia*, *50*(7), 1609-1620.
- Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody: evidence from ERPs. *Brain research*, *1217*, 171-178.
- Paulmann, S., Pell, M. D., & Kotz, S. A. (2009). Comparative processing of emotional prosody and semantics following basal ganglia infarcts: ERP evidence of selective impairments for disgust and fear. *Brain research*, *1295*, 159-169.
- Paulmann, S., Seifert, S., & Kotz, S. A. (2010). Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. *Social neuroscience*, *5*(1), 59-75.
- Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: influence of functional and structural attributes of prosodic contours. *Neuropsychologia*, *36*(8), 701-715.
- Pell, M. D., & Baum, S. R. (1997). The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain and Language*, *57*(1), 80-99.
- Picton, T. W., D. L. Woods, J. Baribeau-Braun, and T. M. G. Healey. 1977. Evoked potential audiometry. *Journal of Otolaryngology*, *6*(2), 90-119.
- Plante, E., Creusere, M., & Sabin, C. (2002). Dissociating sentential prosody from sentence processing: activation interacts with task demands. *NeuroImage*, *17*(1), 401-410.
- Ross, E. D. (1981). The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology*, *38*(9), 561-569.
- Ross, E. D., & Mesulam, M. M. (1979). Dominant language functions of the right hemisphere? Prosody and emotional gesturing. *Archives of Neurology*, *36*(3), 144.
- Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain and language*, *104*(1), 51-74.
- Rymarczyk, K., & Grabowska, A. (2007). Sex differences in brain control of prosody. *Neuropsychologia*, *45*(5), 921-930.
- Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *Neuroimage*, *28*(4), 848-858.
- Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of cognitive neuroscience*, *15*(8), 1135-1148.
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in cognitive sciences*, *10*(1), 24-30.

- Schirmer, A., Chen, C. B., Ching, A., Tan, L., & Hong, R. Y. (2013). Vocal emotions influence verbal memory: Neural correlates and interindividual differences. *Cognitive, Affective, & Behavioral Neuroscience, 13*(1), 80-93.
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research, 14*(2), 228-233.
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: Sex differences revisited. *Cognitive Brain Research, 24*(3), 442-452.
- Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport, 16*(6), 635-639.
- Schlanger, B. B., Schlanger, P., & Gerstman, L. J. (1976). The perception of emotionally toned sentences by right hemisphere-damaged and aphasic subjects. *Brain and Language, 3*(3), 396-403.
- Seddoh, S. A. (2002). How discrete or independent are "affective prosody" and "linguistic prosody"? *Aphasiology, 16*(7), 683-692.
- Sidtis, J. J., & Van Lancker Sidtis, D. (2003, May). A neurobehavioral approach to dysprosody. In *Seminars in Speech and Language* (Vol. 24, No. 2, pp. 93-106). THEIME MEDICAL PUBLISHERS INC.
- Starkstein, S. E., Federoff, J. P., Price, T. R., Leiguarda, R. C., & Robinson, R. G. (1994). Neuropsychological and neuroradiologic correlates of emotional prosody comprehension. *Neurology, 44*(3 Part 1), 515-515.
- Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research, 30*(3), 267-295.
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature neuroscience, 2*(2), 191-196.
- Tracy, D. K., Ho, D. K., O'Daly, O., Michalopoulou, P., Lloyd, L. C., Dimond, E., Matsumoto, K., & Shergill, S. S. (2011). It's not what you say but the way that you say it: an fMRI study of differential lexical and non-lexical prosodic pitch processing. *BMC neuroscience, 12*(1), 128.
- Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Research on Language & Social interaction, 13*(2), 201-277.
- 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, Language and Hearing Research, 35*(5), 963.
- Walker, J. P., Daigle, T., & Buzzard, M. (2002). Hemispheric specialisation in processing prosodic structures: Revisited. *Aphasiology, 16*(12), 1155-1172.
- Wambacq, I. J., & Jerger, J. F. (2004). Processing of affective prosody and lexical-semantics in spoken utterances as differentiated by event-related potentials. *Cognitive Brain Research, 20*(3), 427-437.

- Weintraub, S., Mesulam, M., & Kramer, L. (1981). Disturbances in prosody: A right-hemisphere contribution to language. *Archives of Neurology*, 38(12), 742.
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody--influence of acoustic parameters and arousal. *Neuroimage*, 39(2), 885-893.
- Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in brain research*, 156, 249-268.
- Wildgruber, D., Ethofer, T., Grandjean, D., & Kreifelts, B. (2009). A cerebral network model of speech prosody comprehension. *International Journal of Speech-Language Pathology*, 11(4), 277-281.
- 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(12), 1384-1389.
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage*, 24(4), 1233-1241.
- 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*.
- 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(13), 3722-3738.