Available online at www.sciencedirect.com
ScienceDirect
Brain and Language 105 (2008) 59–69
Brain and Language
www.elsevier.com/locate/b&l

An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context

Silke Paulmann [a,*], Sonja A. Kotz [a,b]

[a] *Max Planck Institute for Human Cognitive and Brain Sciences, Neuropsychology of Language, P.O. Box 500 355, 04303 Leipzig, Germany*
[b] *Day Care Clinic of Cognitive Neurology, University of Leipzig, Germany*

Accepted 12 November 2007
Available online 3 January 2008

Abstract

Previous evidence supports differential event-related brain potential (ERP) responses for emotional prosodic processing and integrative emotional prosodic/semantic processing. While latter process elicits a negativity similar to the well-known N400 component, transitions in emotional prosodic processing elicit a positivity. To further substantiate this evidence, the current investigation utilized lexical-sentences and sentences without lexical content (pseudo-sentences) spoken in six basic emotions by a female and a male speaker. Results indicate that emotional prosodic expectancy violations elicit a right-lateralized positive-going ERP component independent of basic emotional prosodies and speaker voice. In addition, expectancy violations of integrative emotional prosody/semantics elicit a negativity with a whole-head distribution. The current results nicely complement previous evidence, and extend the results by showing the respective effects for a wider range of emotional prosodies independent of lexical content and speaker voice.

Keywords: Emotional prosody; Emotional semantics; ERPs; Integration

1. Introduction

When we speak we can use a variety of emotional intonations to give a specific meaning to a verbal expression. Dependent on the rise and fall pattern of the voice the interpretation of an utterance can change. Thus, emotional speech relies on psychoacoustic parameters such as fundamental frequency (F_0) and intensity, or loudness. Together with speech rate, or duration, and rhythm, these parameters constitute emotional prosody. Banse and Scherer (1996) defined acoustical profiles in vocal emotion expression with the help of several acoustic parameters. One of their main findings was that each emotion seems to have its own acoustic profile. For example, the vocalization of anger reveals a higher F_0 than the vocalization of sadness. Furthermore, intensity measures reveal louder vocalizations of happy than of sad utterances. To properly communicate emotional meaning, emotional prosody needs to be integrated with emotional semantics (Kotz & Paulmann, 2007) next to other non-verbal functions such as mimicry and gestures (e.g., Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000). Even though such a cognitive process is crucial in every day communication, little is known about the underlying mechanism and the time-course of such an integration process. Therefore, the main aim pursued in the current research was to further substantiate the relative contribution of emotional channels, and the time-course of both emotional prosody and of emotional prosody with emotional semantics.

One general finding in the emotion literature is that emotional stimuli seem to be processed differently from neutral stimuli. One explanation that has been put forward for this processing difference is the evolutionary significance of emotions, i.e., emotional stimuli can lead to prior-

* Corresponding author. Fax: +1 514 398 8123.
E-mail address: Silke.Paulmann@mail.mcgill.ca (S. Paulmann).

0093-934X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2007.11.005

itized processing strategies (see also e.g., Schupp, Junghöfer, Weike, & Hamm, 2004). These prioritized processing strategies may enforce attentional orienting or faster processing of emotional stimuli. Indeed, there is accumulating electrophysiological evidence that emotional differentiation occurs as early as 200 ms after stimulus onset during both *visual* emotional word processing (e.g., Begleiter, Porjesz, & Garozzo, 1979; Schapkin, Gusev, & Kuhl, 2000), and *auditory* emotional sentence processing (e.g., Paulmann & Kotz, in press). More specifically, we have shown that this first differentiation occurs independent of attention, i.e., under implicit processing demands, for both lexical- (Paulmann & Kotz, in press) and non-lexical vocal emotional stimuli (Paulmann, 2006).

However, not only the processing of vocal emotions *per se*, but also the integration of emotional channels seems to occur outside the focus of attention (e.g., Kotz & Paulmann, 2007; Pourtois et al., 2000). For instance, Pourtouis and colleagues investigated the integration of emotional auditory and visual stimuli, and report a larger N100 amplitude for congruent than incongruent visual and auditory emotional stimuli pairs (Pourtois et al., 2000). Moreover, Bostanov and Kotchoubey (2004) studied affective prosodic recognition in a context violation paradigm using emotional exclamations (e.g., "Wow", "Oooh", etc.). They reported a N300 to contextually incongruous exclamation, and interpret this component as an indicator of expectancy violation comparable to the well-known N400. If early negativities reported in the literature (e.g., Bostanov & Kotchoubey, 2004; Kotz & Paulmann, 2007) are indeed a form of the classical N400, this would provide evidence that emotional context may speed up speech processing, leading to an earlier onset of well established event-related brain potential (ERP) components (e.g., Bostanov & Kotchoubey, 2004; Kotz & Paulmann, 2007). Last, in a recent study by Wambacq and Jerger (2004), a standard oddball paradigm was applied to investigate emotional prosodic and lexical-semantic information processing of single words. The authors report a larger N400 amplitude in response to semantic than to prosodic stimuli. In contrast, they found a larger P3a amplitude for prosodic than for semantic stimuli. Interestingly, for stimuli carrying both emotional prosodic and semantic information a reduced N400 and an increased P3a was observed. Arguing along similar lines as Besson and colleagues (Besson, Magne, & Schön, 2002), Wambacq and Jerger (2004) proposed that even irrelevant semantics cannot be ignored. In short, there is accumulating evidence that processing vocal emotions (with or without the integration of further emotional information, such as semantics) is a very rapid and non-attentional process. Moreover, some emotional channels may be more important than others during emotional processing, e.g., a predominant processing of emotional semantic over emotional prosodic information has been reported.

In one of our previous studies (Kotz & Paulmann, 2007), we investigated the integrative time-course of emotional prosody with neutral semantics and of emotional prosody with emotional semantics by means of a cross-splicing procedure. Cross-splicing offers a unique possibility to investigate the temporal dynamics of integrative processes as it allows for temporal synchronization of information. Due to the time-locking of critical deviation points in an on-going speech signal with respect to two experimental conditions, the cross-splicing technique allows to directly compare the respective brain responses of two critical conditions. In particular, this technique allows manipulating expectations specific to emotional prosody or to emotional prosody and emotional semantics (see Section 2 for detailed explanation), the two emotional information channels investigated in the present study. Based on the observation that semantically (e.g., Van Petten & Kutas, 1988) and emotionally (e.g., Schirmer, Kotz, & Friederici, 2002; Schirmer, Kotz, & Friederici, 2005) mismatching context information leads to integration difficulties of mismatching information, we argue the following. If expectancy deviation leads to increased integration difficulty, such integration difficulty should result in different event-related potential (ERP) modulations as a function of information specific expectancy deviation. Thus, emotional prosodic expectancy deviation should elicit a different ERP response than expectancy deviation in context of emotional prosody and emotional semantics. While emotional prosodic integration is referred to acoustic correlates such as perceived pitch, duration, and intensity, and their online integration in a speech stream, combined emotional prosodic/semantic integration requires the integration of acoustical information and lexical information. Our previous data clearly revealed qualitatively different ERP components elicited as a function of integration during emotional prosodic processing and combined emotional prosodic/semantic processing. More specifically, we showed that expectancy violations of emotional (angry and happy) prosody elicited a right-lateralized positivity, while combined emotional prosodic/semantic (angry and happy) expectancy violations elicited an early negativity in the ERP (Kotz & Paulmann, 2007). These results are also in line with studies that showed differential processing for linguistic prosodic processing and combined prosodic/semantic processing (Astésano, Besson, & Alter, 2004).

The present study follows up our previous work by extending these different processing dynamics in three dimensions. First, the minimal use of different emotional prosodies (i.e., angry, happy) was extended. Building on the assumption that human emotions engage distinct neural networks (e.g., Adolphs, 2002) that relate to emotion-specific processing mechanisms, we investigated six basic emotional prosodies (anger, disgust, fear, happiness, pleasant surprise, and sadness), even though some previous evidence argues for emotion-independent integration of emotional prosody and semantics (e.g., Kotz & Paulmann, 2007; Schirmer et al., 2002; Schirmer et al., 2005). Thus, one critical research question at stake was to investigate if the integration of emotional prosody and semantics is truly emotion-independent or not. Second, we included dif-