The Influence of Emotional Content on Event-Related Brain Potentials during Spoken Word Processing

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsprogramm Behavior and Cognition

der Georg-August University School of Science (GAUSS)

vorgelegt von

Annika Graß

aus Gehrden (Hannover)

Göttingen, 2016

Betreuungsausschuss:

Prof. Dr. Annekathrin Schacht, NWG Experimentelle Psycholinguistik, CRC Textstrukturen, Universität Göttingen

Dr. Igor Kagan, Decision and Awareness Group, Kognitive Neurowissenschaften, Deutsches Primatenzentrum Göttingen

Prof. Dr. Julia Fischer, Kognitive Ethologie, Deutsches Primatenzentrum Göttingen

Mitglieder der Prüfungskommission

Referentin: Prof. Dr. Annekathrin Schacht, NWG Experimentelle Psycholinguistik, CRC Textstrukturen, Universität Göttingen

Korreferent: Dr. Igor Kagan, Decision and Awareness Group, Kognitive Neurowissenschaften, Deutsches Primatenzentrum Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Julia Fischer, Kognitive Ethologie, Deutsches Primatenzentrum Göttingen

Oliver Schülke, Behavioral Ecology, Johann-Friedrich-Blumenbach-Institute for Zoology & Anthropology, Universität Göttingen

Kurt Hammerschmidt, Kognitive Ethologie, Deutsches Primatenzentrum Göttingen

Uwe Mattler, Experimentelle Psychologie, Institut für Psychologie, Universität Göttingen

Tag der mündlichen Prüfung: 18.10.2016

Table of Contents

I. Summary

In our everyday lives, language is an indispensable tool for communication and for the establishment and preservation of social interactions. Language can be divided in two different modalities, namely the auditory and visual modality. The auditory modality comprises spoken language, whereas the visual modality is composed of the written part of language. Even though a day without speaking is probably inconceivable for most of us, previous research has neglected the investigation of effects of emotional meaning on spoken word processing, as opposed to processing of written words. For written word processing, several studies elaborately investigated effects of emotional meaning on event-related brain potentials (ERPs). In contrast to this, emotional content in spoken word processing was investigated only very occasionally and mostly either in its interaction with emotional prosody or focused merely on the existence of a specific component. Therefore, it remains an open question how and at which stages emotional content of spoken words affects event-related brain potentials regardless of emotional prosody and whether it shows similarities to the processing of written emotional words.

In this thesis, I investigate the processing of single spoken words with positive, neutral and negative content, with the objective of understanding whether emotional content of spoken words leads to emotion effects in ERPs and if those are comparable to those shown for written words.

In the first study of this dissertation, spoken words of emotional and neutral content were presented to participants at two different volume levels to elucidate possible interactions of emotion effects with bottom-up attention effects driven by stimulus size. For visual stimuli of emotional content as pictures and written words, stimulus size has been shown to increase emotion-related ERP effects, for example at level of the early posterior negativity (EPN). It was investigated whether this augmented relevance of larger visual stimuli might be transferred to the auditory modality. Negative emotional content leads to an

increased frontal positivity and parieto-occipital negativity between 370 and 530 ms. This component reveals resemblance to the visual EPN, however, the negativity expands further towards central scalp areas. Therefore, the question arises whether this component might reflect an auditory counterpart of the visual EPN. Importantly, no interaction of this emotion-related ERP component with volume level is revealed. The following aspects, if comparing them to the visual modality, point towards a broader difference between visual and auditory language processing: The missing interaction of stimulus size and emotion effects, differences in topographies of the emotion effects and the different latencies compared to the visual EPN.

The second part of this thesis aims at a more direct comparison of emotion effects in the visual and auditory modality. For this purpose, a second study was conducted, in which the same words were presented visually and auditorily to the participants. Spoken words were either produced by a computer-voice (Experiment 1) or a human speaker (Experiment 2). This study was designed in order to investigate the existence of an "auditory EPN" and its boundary conditions. In addition, it was investigated whether the higher social relevance of a human voice augments the emotion effects. In both experiments, emotion effects are evident. For written words, effects are evident between 230 and 400 ms, in the early posterior negativity (EPN), for spoken words between 460 and 510 ms. Interestingly, when considering the scalp distribution of the ERP differences between emotional and neutral auditory words, the effect shows even higher similarity to the visual EPN than in the first part of this thesis. Source localization revealed comparable neural generators in the superior parietal lobule (SPL) and inferior parietal lobule (IPL) in both the visual and auditory EPN time window. The findings indicate similarities in the processing of emotional content across modalities that – at least partly – rely on the same neural system. However, these similarities are surprising since the visual EPN is assumed to reflect enhanced sensory encoding in visual areas. The emotion effects revealed in the studies described above differed in terms of latencies, topographies and the valence that elicits the effect (positive or negative).

Therefore, in the last part of this thesis, I investigate potential causes for these differences. Sex differences at scalp topography level are revealed, however, they can not explain the reported differences between the studies. It is hypothesized that both studies reveal the same *auditory emotion-related component* (*AEC*) in a comparable time frame (Study 1: 477-530; Study 2: 464-515 ms), which was preceded by an earlier emotion effect (371-477 ms) with a N400-like scalp distribution in Study 1. Even though no interactions of emotional content and volume level are revealed, presumably volume level manipulation in the first study changed the context of the experiment, which caused the additional effect.

Even though no verifiable cause for the described differences in emotion effects could be revealed, I was able to show the existence of an auditory emotion-related component that is elicited by emotional (compared to neutral) content during spoken word processing. This component is reflected in an anterior positivity and posterior negativity around 460 to 520 ms after word onset. It is invariantly occurring, unaffected by the social significance of the speaker's voice or by a volume level manipulation. Concerning a comparison of the underlying neural network during the processing of content in spoken and written words, it can be concluded that the processing activates brain areas which are at least partly shared in the SPL and IPL. Even though the scalp distribution of the AEC reveals high similarity to the visual EPN, it is not assumed that this effect reflects an auditory counterpart. This conclusion is drawn first on the fact that the typical EPN-distribution is only revealed when calculating the difference waves of emotional and neutral stimuli. The resulting posterior negativity reflects enhanced activation in visual areas to emotional stimuli. The analysis of the underlying neural generators for the difference between auditory emotional and neutral stimuli do not show significant results. However, underlying topographies of the separated emotion categories reveal that the similarity at the level of difference waves resulted from entirely different scalp distributions.

Future research has to control stimulus material more strictly in terms of word length or recognition point in order to reduce the temporal jitter in the data and determine the neural generators of the auditory emotion-related component.

II. Zusammenfassung

In unserem alltäglichen Leben ist Sprache ein unerlässliches Mittel für Kommunikation und die Umsetzung sozialer Interaktionen. Sprache kann in zwei verschiedene Modalitäten unterteilt werden, in die auditorische und die visuelle Modalität. Die auditorische Modalität umfasst gesprochene Sprache, wohingegen die visuelle Modalität vom geschriebenen Teil der Sprache gebildet wird. Auch wenn ein Tag ohne Sprechen für die meisten von uns unvorstellbar ist, hat die bisherige Forschung die Untersuchung von Effekten bei der Verarbeitung von emotionalem Bedeutungsinhalt in gesprochener Sprache, im Gegensatz zu der Verarbeitung von geschriebener Sprache, vernachlässigt. Die Verarbeitung des emotionalen Bedeutungsinhalts von geschriebenen Wörtern hat eine Vielzahl von Studien mit Hilfe von ereigniskorrelierten Potentialen (EKPs) ausführlich untersucht. Im Gegensatz dazu wurde der emotionale Bedeutungsinhalt bei der Verarbeitung von gesprochener Sprache nur gelegentlich und meist entweder in seiner Interaktion mit emotionaler Prosodie oder fokussiert auf die Existenz einer spezifischen EKP Komponente untersucht. Daher bleibt die Frage offen, wie und an welchen Verarbeitungsschritten der emotionale Inhalt gesprochener Sprache ereigniskorrelierte Potentiale beeinflusst, unabhängig von emotionaler Prosodie und der Frage, ob Gemeinsamkeiten mit der Verarbeitung von geschriebenen emotionalen Wörtern bestehen.

In dieser Dissertation untersuche ich die Verarbeitung von gesprochenen Einzelwörtern mit positivem, neutralem und negativem Inhalt, mit der erkenntnisleitenden Fragestellung, ob der emotionale Inhalt von gesprochenen Wörtern Emotionseffekte in EKPs hervorruft und ob diese vergleichbar sind zu denen, die für geschriebene Wörter gezeigt wurden.

In der ersten dieser Dissertation zugrundeliegenden Studie wurden gesprochene Wörter mit emotionalem und neutralem Inhalt den Versuchspersonen in zwei verschiedenen Lautstärken präsentiert, um mögliche Interaktionen mit *bottom-up* Aufmerksamkeitseffekten, geleitet durch die Größe des Stimulus, zu erklären. Für visuelle Stimuli mit emotionalem Inhalt, wie Bilder oder geschriebene Wörter, hat die Größe des Stimulus erhöhte emotionsbedingte EKPs hervorgerufen, zum Beispiel auf der Ebene der *early posterior negativity* (EPN). Es wurde untersucht, ob diese erhöhte Relevanz von größeren visuellen Stimuli auf die auditorische Modalität übertragbar sein könnte. Negativer emotionaler Bedeutungsinhalt führt zu einer erhöhten frontalen Positivierung und einer parieto-okzipitalen Negativierung zwischen 370 und 530 Millisekunden. Diese Komponente zeigt Ähnlichkeit mit der visuellen EPN, obwohl sich die Negativierung zu zentraleren Arealen der Kopfoberfläche ausweitet. Daher stellt sich die Frage, ob diese Komponente das auditorische Pendant zu einer visuellen EPN darstellen könnte. Entscheidend ist hier, dass keine Interaktion dieser emotionsbedingten EKP Komponente mit dem Lautstärkefaktor beobachtet werden kann. Die folgenden Vergleichsaspekte deuten auf umfassendere Unterschiede zwischen visueller und auditorischer Sprachverarbeitung hin: die fehlende Interaktion zwischen der Größe des Stimulus und der Emotionseffekte, die Unterschiede in den Topographien der Emotionseffekte sowie unterschiedliche Latenzen verglichen zu der visuellen EPN.

Der zweite Teil dieser Dissertation ist auf einen direkteren Vergleich von Emotionseffekten in der visuellen und auditorischen Modalität ausgerichtet. Zu diesem Zweck wurde eine zweite Studie durchgeführt, in der Versuchspersonen dieselben Wörter in geschriebener und gesprochener Modalität präsentiert bekamen. Die gesprochenen Wörter wurden dabei sowohl von einer Computerstimme (Experiment 1) als auch von einer menschlichen Stimme (Experiment 2) produziert. Diese Studie wurde konzipiert, um die Existenz einer „auditorischen EPN“ und ihre Randbedingungen zu untersuchen. Darüber hinaus sollte die These überprüft werden, ob die höhere soziale Relevanz einer menschlichen Stimme die Emotionseffekte vergrößert. In beiden Experimenten zeigen sich Emotionseffekte. Für geschriebene Wörter zwischen 230 und 400 Millisekunden, im Zeitbereich der *early posterior negativity*, für gesprochene Wörter zwischen 460 und 510 Millisekunden. Wenn man die Verteilung der EKP Differenzen zwischen emotionalen und neutralen auditorischen

Wörtern berücksichtigt, zeigen die Effekte interessanterweise sogar eine größere Ähnlichkeit mit der visuellen EPN als die Ergebnisse des ersten Teils dieser Dissertation. Eine Quellenlokalisierung ergab vergleichbare neuronale Generatoren im superioren parietalen Lobus (SPL) und im inferioren temporalen Lobus (IPL), sowohl im visuellen als auch im „auditorischen EPN" Zeitfenster. Diese Befunde deuten auf Gemeinsamkeiten in der Verarbeitung emotionaler Inhalte über die Modalitäten hinweg hin, die – zumindest teilweise – durch das gleiche neuronale System gestützt werden. Trotzdem erscheinen diese Gemeinsamkeiten überraschend, da für die visuelle EPN angenommen wird, dass sie eine verstärkte sensorische Enkodierung für emotionale Stimuli in visuellen Arealen abbildet. Die oben beschriebenen und in diesen Studien gezeigten Emotionseffekte unterscheiden sich bezüglich ihrer Latenzen, Topographien und der Valenz, welche den Effekt hervorruft (positiv oder negativ).

Im letzten Teil der Dissertation wurden daher systematisch Unterschiede zwischen den Studien untersucht um potenzielle Ursachen für die oben aufgeführten Unterschiede in den Emotionseffekten bestimmen zu können. Es zeigen sich Geschlechterunterschiede in den Topographien in Studie 2, die jedoch nicht die gefundenen Unterscheide in den Emotionseffekten zwischen den beiden Studien erklären können. Es wird angenommen, dass beide Studien die gleiche auditorische emotions-bedingte Komponente (AEK) in einem vergleichbaren Zeitfenster (Studie 1: 477-530 ms; Studie 2: 464-515 ms) hervorrufen, welcher in der ersten Studie eine N400-ähnlichen Verteilung vorausgegangen ist. Obwohl keine Interaktionen zwischen emotionalem Inhalt und Lautstärke aufgezeigt werden können, gehe ich davon aus, dass die Manipulation der Lautstärke in der ersten Studie den Kontext des Experiments verändert, und so den früheren Effekt ausgelöst hat.

Auch wenn keine verifizierbaren Ursachen für die beschriebenen Unterschiede zwischen den Emotionseffekten aufgezeigt werden konnten, ist es mir mit dieser Dissertation gelungen, die Existenz einer auditorischen emotions-bedingten Komponente zu zeigen, die durch emotionalen (in Vergleich zu neutralem) Inhalt während der Verarbeitung von gesprochener Sprache hervorgerufen wird. Diese Komponente spiegelt sich in einer anterioren Positivierung und einer posterioren Negativierung zwischen 460 und 520 Millisekunden nach Wortbeginn wider. Diese zeigt sich gleichbleibend, unabhängig von der sozialen Signifikanz der Stimme des Sprechers oder der Manipulation der Lautstärke. Bezüglich eines Vergleich des zugrundeliegenden neuronalen Netzwerkes während der Verarbeitung des Inhalts von gesprochenen und geschriebenen Wörtern, kann man annehmen, dass die Verarbeitung Hirnareale aktiviert, die zumindest teilweise im SPL und IPL liegen. Obwohl die Verteilung der AEK eine hohe Ähnlichkeit zur visuellen EPN aufzeigt, kann man nicht annehmen, dass dieser Effekt ein auditorisches Pendant darstellt. Diese Schlussfolgerung beruht darauf, dass sich eine typische EPN-Verteilung nur bei der Berechnung der Differenzkurven von emotionalen und neutralen Stimuli zeigt. Die daraus resultierende posteriore Negativierung spiegelt eine erhöhte Aktivierung von visuellen Arealen - hervorgerufen durch emotionale Stimuli - wider. Die Analyse der zugrundeliegenden neuronalen Generatoren für den Unterschied zwischen auditorischen emotionalen und neutralen Stimuli liefert keine signifikanten Ergebnisse. Trotzdem zeigen die zugrundeliegenden Topographien der einzelnen Emotionskategorien, dass die Gemeinsamkeit auf der Ebene der Differenzkurven aus völlig unterschiedlichen Verteilungen resultiert.

Zukünftige Forschung müsste das auditorische Stimulusmaterial bezüglich der Wortlänge oder des Worterkennungspunktes strikter kontrollieren, um den zeitlichen Jitter in den Daten zu reduzieren und somit die neuronalen Generatoren einer auditorischen emotions-bedingten Komponente besser bestimmen zu können.

1 Introduction

Humans are highly social beings. Connecting to others is the basis not only to our survival, but also to our well-being and the success in life. On of the crucial elements and tools for the establishment and the preservation of social relationships in humans is language. Language comprises two modalities, namely a written and a spoken modality. In general, it holds the ability to convey objective information, to find an agreement or to express and share one's inner state and emotions. Additionally, language has the power to trigger emotions in the recipient in communicative situations or during reading. In fact, emotionally salient stimuli are assumed to be particularly relevant to human beings. This preference was suggested to have a neural foundation: Since emotional stimuli seem to indicate beneficial consequences or even appear to be crucial for survival, these stimuli get a prioritized access to the brain's resources (Pourtois et al., 2013). Emotional language, both written and spoken, constitutes a special class of emotional salient stimuli: language is an arbitrary system that requires the translation of symbolic stimuli for the obtainment of emotional salience and content per se. Thus, linguistic stimuli are considered to be of less biological relevance as compared to pictures of emotional objects and scenes or faces which are considered to be evolutionary prepared. Nevertheless, it was shown in several studies that linguistic stimuli not only hold the ability to elicit emotions in the reader or listener, but also show typical emotion effects in behavioral measures (Schacht and Sommer, 2009a, 2009b; Bayer et al., 2011) and brain correlates (e.g. Kissler et al., 2007; Herbert et al., 2009; Palazova et al., 2013). For the written modality of language, the boundary conditions of emotional facilitation have been intensively investigated in recent years. However, previous research has largely neglected the study of emotion effects in spoken word processing, although spoken language plays a major role in human communication in our everyday lives. Compared to the written modality, spoken utterances comprise two distinct communication- and information channels: Apart from the semantics (as comparable with the written modality), also the tone of a speaker's voice matters and gives us additional information, i.e. the prosody. The majority of existing studies investigating the impact of emotion on auditory processing focuses on emotional prosody or the interplay of emotional prosody and semantics.

The main aim of the present work is to contribute to the question how emotional content of spoken words would affect event-related brain potentials regardless of emotional prosody and to reveal similarities with or differences to the written modality. For this purpose, in two studies comprising three experiments, we investigated the processing of single, spoken nouns produced with a neutral prosody. In the first study (chapter 2), it was investigated whether emotional content of spoken words leads to emotion effects in ERPs comparable to the written modality. Our particular interest focused on whether an interaction with bottom-up attention effects driven by stimulus size applies similarly to the spoken modality, in order to elucidate whether these effects have a comparable functional locus.

The first study revealed only little indication for emotion effects in the auditory modality of language that are similar to those in the visual modality. Therefore, the second study (chapter 3) aimed at a direct comparison of emotion effects in both modalities. Consequently, we conducted two experiments in which the same words were either presented in their written or auditory form, the latter produced by a computer-generated voice (Experiment 1) or a human voice (Experiment 2). This study focused on the question whether emotion effects evoked by spoken words show similarity to the emotion effects evoked by the same written words. By means of this design, we tried to determine whether a comparable system for processing emotional content in both modalities of language can nevertheless be assumed, opposing the differences indicated by the first study. Furthermore, it was of interest whether the emotion effects in Study 1 were reproducible in Study 2. The different voices were used to investigate if there might be effects of subtle, content-related prosody in the human voice or if emotion effects might even be potentiated by the naturalness of the human voice, which might result in higher social relevance.

1.1 **Theoretical and Empirical Background**

Emotionally salient stimuli are assumed to be of special relevance to human beings and do easily attract the attention of an observer (Pourtois et al., 2004). This processing advantage is based on the intrinsic motivational relevance of emotional information for the organism that leads to binding of attention and processing resources. Emotional stimuli, both positive and negative, might bear information which is crucial for survival and thus support the organism to quickly initiate appropriate approach or avoidance behaviors in response to salient cues (Lang, 1995). Therefore, stimuli of emotional content receive prioritized and rapid access to the brain's processing resources (Pourtois et al., 2013). On the behavioral level, this preference leads to a faster detection and higher accuracy (Schacht and Sommer, 2009b, 2009a; Bayer et al., 2011) as well as greater influence on task-relevant behaviors (Vuilleumier and Driver, 2007; Mitchell et al., 2008). In the visual modality, these effects of emotion are quite well investigated and assumed to be caused by enhanced activation in the visual cortex (Lane et al., 1999; Herrmann et al., 2008; Alpers et al., 2009). They are considered to be mediated through re-entrant projections from emotion-related brain structures, in particular the amygdala, to the visual cortex (Isenberg et al., 1999; Tabert et al., 2001; Nakic et al., 2006; Sabatinelli et al., 2009; Herbert et al., 2011b). By now, there is some evidence that the involvement of the amygdala in the detection of emotionally salient stimuli is not limited to the visual domain, but also applies to the auditory modality (Johnstone et al., 2006; Fecteau et al., 2007; Anders et al., 2008). Although less research has been conducted in the auditory domain, studies demonstrate a rapid processing of emotional aspects in auditory stimuli as well (Goydke et al., 2004; Sauter and Eimer, 2010). In addition, enhanced activity triggered by emotional compared to neutral stimuli has been observed in the auditory cortex across different techniques, including fMRI (Wiethoff et al., 2008) and fNIRS (Plichta et al., 2011). Together, this evidence indicates that when processing visual and auditory information, emotion might affect similar processing stages, namely by boosting activity in modality-specific sensory cortices (auditory and visual cortex, respectively). Due to the distinguished role of language in social communication and interactions, it seems conceivable to assume that these similarities in the processing of visual and auditory emotional information would also be evident in the processing of written and spoken emotional language. Thus, the detection of emotional content in both written and spoken language is likely to take place in the same areas of the brain (like the amygdala or the parietal lobule), but the re-entrant projections would necessarily target modality-specific perceptual areas, that is the visual and auditory cortices.

Language, compared to pictures of emotional objects and scenes, emotional sounds and facial expression of emotion, constitute a special condition of emotional stimuli. The emotional connotation of linguistic stimuli is symbolic, arbitrary and acquired by learning processes. Therefore, linguistic emotional material is less evolutionary prepared as compared to pictorial stimuli or affective sounds (e.g. a baby's crying, a burning fire, laughter). However, considering the tight link between emotion and language and the important role of language in general, it is not surprising that the processing advantage for emotional information shown in other stimulus domains is also evident in language. The arbitrariness of language applies to both, written and spoken language; however, the modalities differ in their degree of abstraction as well as in their point of phylogenetic and ontogenetic acquisition. Learned within the first years after birth without explicit grammar instructions, spoken language is assumed to have an innate system for grammatical rules and being learned with accomplished effort and repetition (Sakai, 2005). In contrast to this, written language is learned later on in the ontogeny and requires instruction and practice. Furthermore, developmental dyslexia, which is partly caused by structural abnormalities and leads to reading but not speaking disabilities, indicates that reading abilities require specific neural mechanisms (for a review see Stein, 2001). All this might give a hint for a differential processing of the written and spoken modality of language in the human brain and that the spoken part might be more natural and relevant.

As a shared characteristic, the processing of language in both modalities has to be flexible. Written words are recognized as words in seconds, invariantly over changes in font, position, case or size (Dehaene et al., 2005). Comparably, speech is highly variable with different acoustic realizations of words for different speakers, different speaking styles and rates. For both auditory and visual word comprehension, a number of basic processes is required: First, an encoding of the input is needed including the identity and order of letters or phonetic elements. Second, a matching of the input to long-term memory representations in the "mental lexicon" has to take place. And third, the best matching candidate must be selected from the tens of thousands of words in the recipient's vocabulary.

For both modalities, there exist numerous theories and models elucidating in which format lexical knowledge is stored and how it is accessed when needed. McClelland and Rumelhart (1981) have proposed the Interactive Activation (IA) Model – one of the first "neural-network" cognitive models for reading. According to the model, perception results from excitatory and inhibitory interactions of detectors for visual features, letters and words. When a word is presented to a reader, the visual input excites detectors for visual features (curved shapes, horizontal and vertical bars etc.) and at the same time inhibits other feature detectors. Those feature detectors will then stimulate or inhibit different letter detectors, which will finally excite or inhibit word detectors. Each activated connection carries a different weight with the target word being more activated than any other word and therefore recognized by a reader.

Obviously, spoken word processing (at least) at the initial stages relies on a different sensory system. Accordingly, a different model is required to explain the mapping of phonemes onto the word lexicon. The first psycholinguistic model of spoken word recognition was the Cohort Model, developed by Marslen-Wilson (1978; 1980). Even though it relies on a different sensory modality, the model shows similarity to the IA model in the visual modality. The Cohort Model consists of three stages: access, selection and integration. During the access stage, the first one or two acoustic-phonetic elements reach the hearer's ear and the mental lexicon activates every possible word that starts with that speech segment (the cohort). During selection, more phonetic elements enter the ear and candidate words that mismatch the signal by more than one single feature are removed from the cohort. During integration, the semantic and syntactic properties of activated words are retrieved and checked for integrability with higher levels. However, contrary to the IA model, the candidate words in the cohort do not actively compete with one another. Solely the presence of other candidate words forms the recognition process.

Those different processing stages necessary for the understanding of language rely on concurrent activations of multiple areas within a distributed neural network. However, the question remains whether the process of visual language processing does converge with auditory language processing at a certain point or whether it relies on completely separated pathways. Indeed, the general view on the neural processing of language relies on distributed but interactive brain areas, assuming modality-specific lexical components accessing a central semantic system (for reviews see Mesulam, 1998; Cabeza and Nyberg, 2000; Martin, 2003). Marinkovic and colleagues (2003) showed that the phonological input route for understanding spoken words and the orthographic route underlying reading were quite distinct during initial processing, but overlapping areas were subsequently activated during stages of semantic and contextual integration. Using anatomically constrained magnetoencephalography (aMEG), combining high-density whole-head MEG with anatomical magnetic resonance imaging, they revealed the activity to spoken words starts in primary auditory regions and spreads anterolaterally to the lateral superior temporal area at 55 ms, spreading to the perisylvian/superior temporal plane (~100 ms) (the ventral or `what` auditory processing stream) and at ~ 250 ms reaching anterior regions of the temporal lobe (AT), the perisylvian area and posterior inferior prefrontal regions bilaterally. That implies a processing of spoken words in auditory cortex at initial stages, followed by voice-specific processing in the superior temporal area bilaterally (Cabeza and Nyberg, 2000) followed by speech selective areas in the superior temporal sulcus of the left hemisphere (Scott et al., 2000). When reading a word, on the contrary, activity spreads forward from the bilateral occipital area along the ventral visual pathway. Activity peaks in the left ventral temporo-occipital area at ~ 170 ms, in the superior temporal sulcus (STS) and the inferolateral temporal area at ~ 230 ms and in the AT at ~ 350 ms, encompassing the left inferior prefrontal cortex (LIPC) and the orbitofrontal cortex bilaterally at ~ 400 ms. The shown transition from modality specific-streams to the access of supramodal networks for semantic access and contextual integration occurrs approximately 230 ms after word onset and is particularly evident for the N400 component. The N400 is a scalp-recorded negativity peaking at ~ 400 ms which is thought to index access to meaning. This finding favors the claim that modulations in N400 amplitude reflect supramodal semantic processes.

Regardless of their difference concerning ontogenetic and phylogenetic development, the aforementioned findings lead to the assumption of a comparable processing of written and spoken words in the brain (at least at later stages). Coming back to the processing of emotions, the question still remains at which stages in this temporal dynamics of word processing emotion content does interact.

Emotion-related activation in modality-specific areas has been shown for the visual domain using event-related brain potentials (ERPs), which provide a useful tool to investigate the temporal dynamics of word processing. Effects of emotional content were reliably shown for the so-called early posterior negativity (EPN) (Junghöfer et al., 2001). This relative negativity over posterior electrode sites and positivity over fronto–central electrode sites occurs between approximately 250-400 ms after stimulus onset and was shown to emotional facial expressions (Junghöfer et al., 2001; Holmes et al., 2008; Recio et al., 2011; Rellecke et al., 2012) and emotional pictures (Schupp et al., 2004; Schupp et al., 2007; Bayer and Schacht, 2014). By now, numerous studies have investigated the influence of emotional meaning on visual word processing and reported robust effects of emotional content on the EPN (Kissler et al., 2007; Kissler et al., 2009; Schacht and Sommer, 2009b, 2009a; Scott et al., 2009; Palazova et al., 2011; Bayer et al., 2012a, 2012b; Opitz and Degner, 2012; Citron et al., 2013;

Palazova et al., 2013). Interestingly, the effects of emotional valence on the EPN component were shown to interact with the perception of proximity in different domains: the presentation of emotional pictures of bigger size (and therefore presumably perceived higher proximity) enhanced emotion effects for the EPN (De Cesarei and Codispoti, 2006). This interaction of stimulus size and emotion was not only shown for emotional pictures with a high biological relevance and evolutionary preparedness but also for written words (Bayer et al., 2012a). The authors suggested that the mechanisms of sensory facilitation were originally based on a biological relevance, important for survival, but might have generalized to written words.

Functionally, the EPN has been linked to enhanced sensory encoding and attention allocation to emotional stimuli. This enhanced activation in the early visual cortex was shown not only by EEG (Schupp et al., 2003; Keil et al., 2005; Junghöfer et al., 2006) but also by fMRI studies (Lang et al., 1998; Alpers et al., 2009). Since preferential processing of emotionally salient stimuli is not limited to the visual domain, but also occurs for other modalities (e.g. sounds, smells), it is expectable to find a similar boost in auditory brain areas during the processing of acoustically presented words with emotional content. Thus, the question arises whether this enhanced activation of auditory areas might be reflected in an auditory EPN-equivalent in ERPs. To date, only two studies have investigated the existence of such an auditory EPN-equivalent. Mittermeier and colleagues (2011) and Jaspers-Fayer and co-workers (2012) compared brain activations across three tasks including a neutral tone presentation, a prosodic emotion task with emotional uttered syllables and a semantic emotion task with words of emotional content. The former study reported an increased negativity for emotional tasks compared to the non-emotional task with a maximum at 170 ms post-stimulus ("EPN170"), whereas the latter showed larger negativities at Pz electrode in the word and syllables task compared to the neutral tones (i.e. around 130 to 150 ms and 250 to 390 ms after stimulus onset). However, these effects were not caused by enhanced activation in auditory areas; instead, an EEG–fMRI single-trial coupling showed that variance of voltage was correlated with activity in medial prefrontal cortex in an early time frame and superior parietal lobule in later time frames. The superior parietal lobule is associated with the control of selective attention and has previously been identified as one of the sources underlying the (visual) EPN (Junghöfer et al., 2010). Accordingly, Jaspers-Fayer and colleagues (2012) hypothesized a common neural generator of visual and auditory EPN components in the superior parietal lobule. However, the authors' arguments for defining this component as an equivalent to the visual EPN remain speculative. Furthermore, at least some of the effects reported for spoken words of emotional content seem rather early for lexical access to have taken place (cf. Pulvermüller and Shtyrov, 2006). Importantly, it has to be taken into account that Jaspers-Fayer et al. (2012) contrasted emotional word processing to the processing of neutral tones, but not to neutral words. Therefore, the effects reported for acoustically presented emotional words cannot be solely ascribed to differences in emotional meaning, but they also reflect very fundamental differences in stimulus processing.

Recently, Rohr and Abdel Rahman (2015) reported effects of emotional content during spoken word processing that, interestingly, were restricted to an experimental situation providing a minimal social context in the form of short videos of the speaker making direct eye contact with the participants. They reported robust ERP effects of emotional content around 250 ms, in form of an enhanced positivity over posterior electrode sites for emotional compared to neutral words. Remarkably, in non-communicative contexts and in their pre-experiment (written and auditory words without a speaker's face present), the authors reported this effect to be reduced or even absent.

To conclude, it still remains unknown whether an equivalent to the visual EPN exists in the auditory modality and whether such an early, emotion-related activation would lead to enhanced activity in auditory areas. Instead, the same sources underlying the visual EPN might also be active during the processing of emotional auditory stimuli, as it has been shown that auditory stimuli in general can activate visual areas (Qin and Yu, 2013; Feng et al., 2014). Furthermore, it remains unclear whether such emotion effects on an auditory EPN would interact with stimulus-triggered attention driven by physical stimulus-features comparable to the interactions shown for visual EPN.

At a later stage of more elaborate, higher-order processing, visually presented emotional words have been demonstrated to elicit enhanced amplitudes in ERPs over centro-parietal regions – the late positive complex (LPC) (Cuthbert et al., 2000; Herbert et al., 2006; Schacht and Sommer, 2009b; Bayer et al., 2012b; Bayer and Schacht, 2014). The LPC component is thought to reflect more sustained elaborate processing of emotional stimuli and is related to the P300 component occurring in the same time range. The LPC was shown to be elicited not only by pictures but by words as well and its amplitude is modulated by the valence of the presented stimuli (Herbert et al., 2006; Schacht and Sommer, 2009b; Liu et al., 2010). However, effects on the LPC by emotional content seem to be rather unstable, highly context- and task-dependent (Fischler and Bradley, 2006; Schacht and Sommer, 2009b, 2009a), and presumably influenced by top-down attention (e.g. Schupp et al., 2007; Kissler et al., 2009; Bayer et al., 2012a). A counterpart of the visual LPC was also revealed in the auditory domain, which was shown to be modulated by emotional content of stimuli, both for spoken words with emotional connotation (Ofek et al., 2013; Hatzidaki et al., 2015) and emotionally uttered words and sentences (Paulmann et al., 2013; Schirmer et al., 2013). These findings indicate domain-general mechanisms to be involved during the processing of emotional visual and auditory stimuli; however, further evidence for this assumption is needed.

A component not primarily linked to emotion, but occurring to linguistic stimuli across both modalities is the afore-mentioned N400 component. The N400 is known as an indicator of semantic processing, which was shown to be modulated by overall expectancy and congruity of (neutral) stimuli in semantic contexts (cf. Kutas and Federmeier, 2011). By now, there is abundant evidence for the occurrence of the N400 in comparable auditory paradigms as semantic priming paradigms (e.g. Perrin and García-Larrea, 2003; Relander et al., 2009; Erlbeck et al., 2014), congruency paradigms (e.g. Schirmer and Kotz, 2003; Paulmann et al., 2009; Diamond and Zhang, 2016) and to sentences with semantic or syntactic violations (e.g.

Hahne and Friederici, 2002; Wicha et al., 2003; Balconi and Pozzoli, 2005; Lück et al., 2006; Erlbeck et al., 2014). N400 effects to emotional stimuli were shown, but only in typical congruency paradigms where semantic and prosodic emotional information was congruent or incongruent. To my knowledge, no study so far revealed modulations of the N400 by emotional semantic content relative to neutral content in auditory paradigms. For visual paradigms, however, emotional information has been shown to affect the N400 (Kanske and Kotz, 2007; Herbert et al., 2008; Holt et al., 2009).

All in all, it remains unclear if auditory word processing is impacted by emotional content in a manner comparable to visual word processing. It is still an open question whether emotion effects on ERPs would appear with comparable scalp topographies and latencies or show a completely different pattern due to the different underlying sensory processing. As discussed above, there is evidence for the existence of auditory counterparts to several visual (emotion-related) ERP components, but they were shown with varying latencies and topographies. Furthermore, no clear reasons for defining them as equivalents to the visual counterparts were given. In this dissertation, I will investigate effects of emotional content on auditory event-related potentials evoked by single spoken words and compare them to the visual modality. By that I aim at answering the question whether the processing of emotion in language takes a comparable neural pathway in the written and spoken modality or whether it relies on different processing systems.

2 Effects of Volume Level and Emotional Content on Spoken Word Processing (Study 1)[1]

Introduction

From an evolutionary perspective, rapid detection of threats or life-sustaining opportunities is important for survival and fast adaptation and explains the outstanding importance of emotional stimuli for humans. The organization of the emotional response systems has been suggested to be founded on two basic motivation systems, an appetitive and a defensive system (Lang et al., 1997; Lang and Bradley, 2010) leading to approach or avoidance reactions, respectively (Paulus and Wentura, 2014, 2016). Reacting fast to a positive stimulus, for instance, might maximize the probability of attaining a rewarding state, whereas emotionally negative stimuli are best dealt with by initiating a rapid response that probably aids survival. Therefore, it seems conceivable that the high importance of emotional content shapes perceptual processing and finally results in appropriate reactions. Next to somatic reactions, this modulation is evident on the behavioral level in better memory performance (Kissler et al., 2007; Kissler et al., 2009; Bayer et al., 2011), faster response latencies (Keil et al., 2005; Schacht and Sommer, 2009a, 2009b; Bayer et al., 2011) and higher accuracy (Schacht and Sommer, 2009b) for emotional compared to neutral stimuli. The preferential processing of emotional stimuli is also evident in electrophysiological correlates as for example in the EPN component.

Emotional valence also seems to interact with the perception of proximity: Positive objects are perceived as closer than negative and neutral ones (Valdés-Conroy et al., 2012) and the effect of proximity on reaction times was shown to be modulated by the valence of an approaching stimulus (De Haan et al., 2016). Codispoti and De Cesarei (2007) investigated physiological changes and subjective ratings of participants in response to emotional pictures

---

[1] Grass, A.; Bayer, M.; Schacht, A. (2016) - Electrophysiological Correlates of Emotional Content and Volume Level in Spoken Word Processing. Frontiers in Human Neuroscience 19, 347-357. doi: 10.3389/fnhum.2016.00326. The original publication was adapted to the structure of this dissertation.

of varying sizes, as an increase in object size seems to be the main characteristic of an approaching object. They revealed an interaction of stimulus size and emotional reactions: Pictures of large size triggered stronger emotional reactions than smaller pictures, consisting of increased amplitudes of skin-conductance responses as well as more pronounced differences in subjective valence and arousal ratings between emotional and neutral pictures. A similar interaction of emotion and stimulus size was shown for the EPN, which started earlier and was more pronounced for larger than for smaller pictures (De Cesarei and Codispoti, 2006). The authors proposed that an increase in image size might lead to enhanced emotional reactions due to the more direct biological relevance of pictorial stimuli. One could argue that the size of the picture reflects the subjective proximity of a perceiver to a given object in reality and thereby influences its biological relevance. An aggressor, for example, is more dangerous the closer it is. Thus governed by the higher motivational relevance, the response to this stimulus should be more pronounced.

Independent of emotional aspects, amplitudes of two early components related to the processing of visual stimuli, namely N1 and P1, are comparably modulated by both objects in near space (Kasai et al., 2003; Valdés-Conroy et al., 2014) and by bigger images (Nakayama and Fujimoto, 2015; Pfabigan et al., 2015), indicating a close link between image size and proximity. Similarly, these early stages of perceptual processing were shown to be impacted by other stimulus features as brightness, contrast and texture appearance (Johannes et al., 1995; Balas and Conlin, 2015; Schettino et al., 2016).

Bayer and co-workers (2012a) investigated whether the interaction of stimulus size and emotion effects generalizes to linguistic materials, namely to isolated words of emotional meaning. If the interaction of image size and emotion existing for pictures is resulting from the higher biological relevance due to its direct resemblances of the object they depict, a similar effect would be unlikely to occur for written words, since they are entirely arbitrary and symbolic. Interestingly, large stimulus size – more precisely font size of written words – led to augmented ERP effects of emotional content in the EPN time frame, showing high similarity to effects reported for affective pictures (De Cesarei and Codispoti, 2006). The authors thus concluded that the mechanism responsible for interactions of emotional and stimulus-triggered attention might not be limited to biologically relevant stimuli, but might also be engaged in processing of symbolic stimuli. Thus, a more general type of stimulus relevance is possibly playing a causal role in the interaction of stimulus size and emotional content. The authors suggested that the mechanisms of sensory facilitation were originally based on a biological, survival-related type of relevance, but might have generalized to written words, probably reflecting the high social significance of language (Bayer et al., 2012a).

As a consequence, the question arises if and how this mechanism would apply to the spoken domain of language, which may play an even more important role in the everyday life of human beings. Given that an approaching object mainly changes in its physical size, the main characteristic of sounds in near versus distant environment are differences in their volume level (e.g. von Bekesy, 1949; Begault, 1991; for a review on auditory distance perception see Zahorik, 1996). Similar to stimulus size in the visual domain, volume level has been shown to modulate early cortical responses to auditory stimuli. An increase in volume increases the N1/P2 peak-to-peak amplitude (Rapin et al., 1966; Beagley and Knight, 1967; Picton et al., 1970; Adler and Adler, 1991; Thaerig et al., 2008). The N1/P2 complex is a cortical auditory evoked potential reflecting auditory processing. However, to the best of my knowledge, it remains unclear whether there exist later effects of volume level on auditory-evoked potentials and if volume level might also interact with emotion effects, as it has been shown for the stimulus size of emotional pictures and written words. A candidate component for a possible interaction of volume level and emotional content would be an auditory EPN, which was proposed to be an equivalent to the visual EPN in the auditory domain (Mittermeier et al., 2011; Jaspers-Fayer et al., 2012). Next to the semantic content of a sentence or word, spoken utterances comprise a second communication channel, namely prosody. The tone and rhythm of a speaker's voice can convey emotion as well and might be more innate than the learned, artificial meaning of words. Using auditory stimuli of varying emotional prosody and content, two studies (Mittermeier et al., 2011; Jaspers-Fayer et al., 2012) demonstrated a negativity occurring in emotional compared to non-emotional paradigms.

Next to the EPN-counterpart in response to auditory emotional stimuli, evidence also suggests the existence of an equivalent to the late positive complex (LPC), which has reliably been shown to reflect sustained elaborate processing of emotional stimuli in the visual modality. An auditory LPC was reported for spoken words with emotional connotation (Mittermeier et al., 2011; Ofek et al., 2013; Hatzidaki et al., 2015) and emotionally uttered words and sentences (Paulmann et al., 2013; Schirmer et al., 2013). Although there is evidence for some similarities between emotion-related ERP effects in the visual and auditory modality, it is noteworthy that these effects show pronounced differences in their temporal dynamics. Furthermore, strong differences in terms of the latency of emotion-related effects can be found *within* the auditory modality: Whereas emotional prosody conveys salience rather immediately and can thus modulate quite early components as for example the P2 (Paulmann and Kotz, 2008; Agrawal et al., 2012; Pinheiro et al., 2013; Schirmer et al., 2013) full semantic information of spoken words, including their emotional content, incrementally develops over time (Bradley and Lang, 2000). Therefore, the time course of effects for emotional meaning is rather difficult to compare to effects for emotional prosody, but also to effects of emotional meaning in the visual modality.

The aim of the present study was to investigate the interplay of volume level and emotion effects for the auditory domain of language, using the stimulus material of Bayer and co-workers (2012a)[2]. Stimuli were spoken in neutral prosody by a trained female speaker and presented in two different volume levels. First, effects of volume on the N1-/P2- complex were expected. Whereas for the written domain of word processing, effects of emotional content on early components as the P1 have been reported (Hofmann et al., 2009; Bayer et al., 2012b; Hinojosa et al., 2015), early emotion effects for the auditory modality were not expected due

---

[2] Two nouns of the original stimulus material of Bayer et al. (2012a) had to be replaced because of their ambiguous phonology.

to the following reasons: First, to our knowledge there is no evidence for impacts of emotional content on early ERP components in the auditory domain, except for natural tone stimuli that were associated with emotion in conditioning paradigms (Bröckelmann et al., 2011) and effects of emotional prosody as reported before. Although explicit ratings as well as autonomous measures indicate high similarities between affective picture and affective sound processing in terms of perceived emotional arousal and valence (Bradley and Lang, 2000; Partala and Surakka, 2003), early ERP modulations to nonlinguistic affective sounds have not yet been reported. Thierry and Roberts (2007) implemented a combination of an oddball paradigm and a one-back matching task, in which neutral sounds were presented at two different volume levels (standard versus deviants), additionally interspersed with unpleasant sounds presented at low volume level (deviants). Importantly, volume differences within the neutral stimuli impacted early ERP components (N1, P2) whereas effects of unpleasantness became evident only after about 300 ms. Second, the study of Bayer and colleagues (2012a) – using the same word stimuli and a highly similar paradigm as we employed in this study – did not show emotion effects at the P1 level in the visual domain. Third, in the present study, ERPs were measured to the words' onsets. Thus, during initial processing stages – as reflected by the auditory N1-P2 complex – only a small amount of (semantic) information is available. This incremental nature of auditory processing of rather complex stimuli as words and sounds might also explain the absence of early effects in the study of Thierry and Roberts (2007).

In accordance with the claim of the existence of an auditory EPN (Mittermeier et al., 2011; Jaspers-Fayer et al., 2012), we expected an emotion-related ERP effect, consisting of an enhanced frontal positivity and posterior negativity (comparable to the visual counterpart). Assuming that this component is a functional equivalent to the visual EPN, the volume level should modulate these emotion effects, similar to interactions reported for emotional pictures and written words. This modulation should be limited to sensory encoding, while no interactions at higher-order processing stages should occur (De Cesarei and Codispoti, 2006; Bayer et al., 2012a). If, in contrast, the mechanism underlying the interplay of stimulus size and emotion is restricted to the visual modality, effects of emotional content and volume level in auditory word processing should be independent.

Material and methods

Participants

Data was collected from 31 female participants. Two data-sets had to be discarded due to excessive ERP-artifacts. The remaining participants had a mean age of 23.7 years (SD = 2.8 years), were all right-handed (Oldfield, 1971), native German speakers, and reported no neurological or psychiatric disorders. Participants reported normal hearing range, which was further ensured by a short, custom-made hearing test administered prior to the experiment in which subjects had to count single tones at different volume levels. Participation was reimbursed with course credit or 8 euro per hour.

Stimuli

Stimuli consisted of 72 German nouns that were of positive, neutral, or negative valence (n = 24 each, for word list see Appendix A). The three emotion categories differed significantly in their valence ratings, $F(2, 69) = 1362.67$, $p \leq .001$ (all rating values were drawn from the Berlin Affective Word List Reloaded (Võ et al., 2009); with lower ratings for negative compared to neutral, $F(1, 46) = 725.7$, $p \leq .001$, and higher ratings for positive compared to both negative, $F(1, 46) = 2446.8$, $p \leq .001$, and neutral words, $F(1, 46) = 727.74$, $p \leq .001$ (for descriptive statistics see Table 2.1). Neutral words were significantly less arousing than positive and negative words, $Fs(1,46) > 99.0$, $ps < .001$ which did not differ from each other, $F(1,46) = 1.68$, $p = .202$. Emotion categories were controlled with regard to imageability, word frequency, and the number of letters and syllables, all $Fs(2,69) \leq 1$.

Words were spoken by a trained female speaker in neutral prosody and were recorded on a PC workstation using Adobe Audition (Adobe Systems Software, Dublin, Ireland). In a first step, mean amplitudes for each word were normalized; the analysis of acoustic parameters was then performed using Praat software (Boersma and Weenik, 2009). Emotion categories did not differ in amplitude, mean F0 (fundamental frequency), F0 variability, F0 peak values, overall duration, and speed per syllable. Stimuli were presented in two sound volumes. Based on a pilot experiment, volume levels were adapted in such a way that stimuli were audible in the low volume condition and not too loud in the high volume condition, in order to prevent participants from startling. The mean amplitudes were 43.0 dB (SD=1.6 dB) in the low volume condition and 56.1 dB (SD=2.5 dB) in the high volume condition, measured by a professional sound level meter (SL-322; ATP Messtechnik GmbH), placed at the approximate position of participants' heads. Maxima in volume level did not exceed 67 dB, minima were above 35 dB; thus all words stimuli were presented within the normal range of human communication (e.g., Schwartz and Krantz, 2016). Importantly, volume levels did not differ as a function of emotion, *F*s < 1, while both volumes significantly differed between the two loudness conditions as intended, *F*(1,138) = 1363.6, *p* < .001 (see Table 2.1).

Table 2.1 Descriptive statistics (Means and Standard Deviations) of linguistic and auditory stimulus word parameters.

Parameter	Positive	Neutral	Negative
Valence	2.1 (0.2)	0.3 (0.2)	-2.0 (0.3)
Arousal	3.3 (0.7)	1.9 (0.2)	3.5 (0.5)
Imageability	5.4 (0.8)	5.6 (0.4)	5.5 (0.6)
Letters	6.3 (1.9)	6.3 (1.2)	6.4 (2.1)
Syllables	2.0 (0.8)	2.0 (0.8)	2.1 (1.0)
Frequency	27.7 (32.0)	24.6 (29.2)	24.8 (20.5)
Duration	682.2 (123.6)	628.5 (99.3)	694.6 (149.3)
F0 Range	61.4 (24.2)	66.0 (14.1)	57.3 (14.9)
Mean F0	207.2 (8.8)	202.4 (6.8)	205.3 (8.5)
Low volume level	43.0. (1.9)	43.0 (1.2)	43.1 (1.7)
High volume level	55.8 (2.5)	56.0 (2.2)	56.1 (2.5)

For all ratings, the ranges are: -3 to +3 (valence), 1–5 (arousal), 1–7 (imageability). Frequency is indicated as occurrence per 1 million words in the CELEX database. Note that these values refer to the written version of our word stimuli (Võ et al., 2009). Mean F0, F0 range, and duration were measured in Praat (Boersma and Weenik, 2009) and are indicated in Hertz and milliseconds, respectively. High and low volume level are given in decibel.

Procedure

The study was approved by the ethics committee of the Department of Psychology at the University of Goettingen, Germany, and was conducted according to the Declaration of Helsinki. Before the beginning of the experiment, participants were acquainted with the experimental procedure and signed informed consent. After preparation of EEG recordings, participants were seated in a sound-attenuated chamber. Participants were facing a computer monitor at a distance of 100 cm while words were presented by two loudspeakers positioned at a distance of 133 cm from the participant's ears. The experiment consisted of four experimental blocks; within each block, each word was presented once. Half of the words per block were randomly presented at high volume and the other half at low volume, resulting in two presentations of each word at each volume level in total. The assignment of words to volume levels changed after each block and the order of this assignment, i.e. whether the first presentation of a word was at high or low volume, was counterbalanced. Participants were instructed to listen attentively to the presented words. A one-back task was employed at random intervals (on average after every 9th trial) in order to ensure that participants were paying attention to the word stimuli during the experimental session. In these test trials, a word was displayed within a green frame on the screen. Participants had to indicate by button press whether this word was identical or different to the one they had heard before. By presenting the words in their written form, semantic processing of the words was ensured since the task could not be performed on the basis of perceptual matching. During the presentation of each spoken word, a fixation cross was presented on the screen, starting 1000 ms prior the word onset and remaining visible for 2000 ms after word onset in order to avoid visual offset potentials. The inter-trial-interval (blank screen) had a length of 1000 ms, resulting in an overall trial length of 4000 ms.

EEG recordings and preprocessing

The EEG was recorded with the Biosemi ActiveTwo (Biosemi, Amsterdam, The Netherlands) system from 64 electrodes mounted in an electrode cap (Easy-Cap, Biosemi). Six additional electrodes were placed at the outer canthi and below both eyes in order to record the electrooculogram; two electrodes were placed at the mastoids. The common mode sense (CMS) active and the driven right leg (DRL) passive electrode were used as reference and ground electrodes, respectively (cf. www.biosemi.com/faq/cms&drl.htm). Electrode offsets were kept below a threshold of +/-20mV. Signals were recorded at a sampling rate of 512 Hz and a bandwidth of 104 Hz. Offline, data was processed with the BrainVision Analyzer (Brain Products GmbH, Munich, Germany). The continuous EEG signal was re-referenced to average reference and segmented into epochs of 1200 ms, starting 200 ms prior to word onset. Blinks were corrected using the Surrogate Multiple Source Eye Correction as implemented in Besa (Brain Electric Source Analysis, MEGIS Software GmbH, Gräfeling, Germany); segments containing artifacts (5.4 %) were rejected (voltage steps larger than 50 µV, 200µV/200ms intervals difference of values, amplitudes exceeding -150 µV/150 µV, and activity smaller than 0.5 µV). The overall number of discarded trials per condition (volume level by emotion) ranged between 0 and 19 and did not differ between conditions, as indicated by a repeated-measures ANOVA, all *F*s < .1. Segments were referred to a 200 ms pre-stimulus baseline and averaged per subject and experimental condition.

Data analysis

Segmentation of ERP amplitudes proceeded according to visual inspection of measures of global field power (Lehmann and Skrandies, 1980) and global map dissimilarity (GMD; Brandeis, 1992). Figure 2.1 depicts GFP contrasted for the factors emotion (positive, negative, neutral) and volume level (low, high), as well as GMD, which was calculated across the six experimental conditions. GFP reflects the overall ERP activity across the scalp at any given moment. GMD reflects the dissimilarity between scalp topographies of adjacent time points and demarcates the borders between periods of relatively stable topographies indicating continued processing within similar brain areas. These transition times were used as the limits of the time segments, for which mean ERP amplitudes were calculated. As becomes obvious from Figure 2.1, GMD peaks were clearly observable at the following time points 0, 30, 80, 130, 265, and 530 ms. In order to allow for more fine-grained analyses of ERPs during the interval of main interest, data was additionally sub-segmented between 265 and 530 ms into five time intervals of equal length (53 ms each). After the last clear segment border, consecutive time frames of 50 ms were analyzed between 530 and 980 ms. Amplitude differences were assessed by repeated measures ANOVAs within these time borders, including the factors emotion (3 – positive, negative, neutral) and volume level (2 - high, low) and electrode (64). Degrees of freedom in ANOVAs were adjusted using Huynh–Feldt corrections. If indicated by significant electrode x emotion, electrode x volume level, or electrode x volume level x emotion interactions in these exploratory analyses, these effects were further tested in region of interests (ROIs) that were defined based on visual inspection of the ERP difference waves within the specific time frames. For post-hoc comparisons, *p*-values were Bonferroni adjusted.

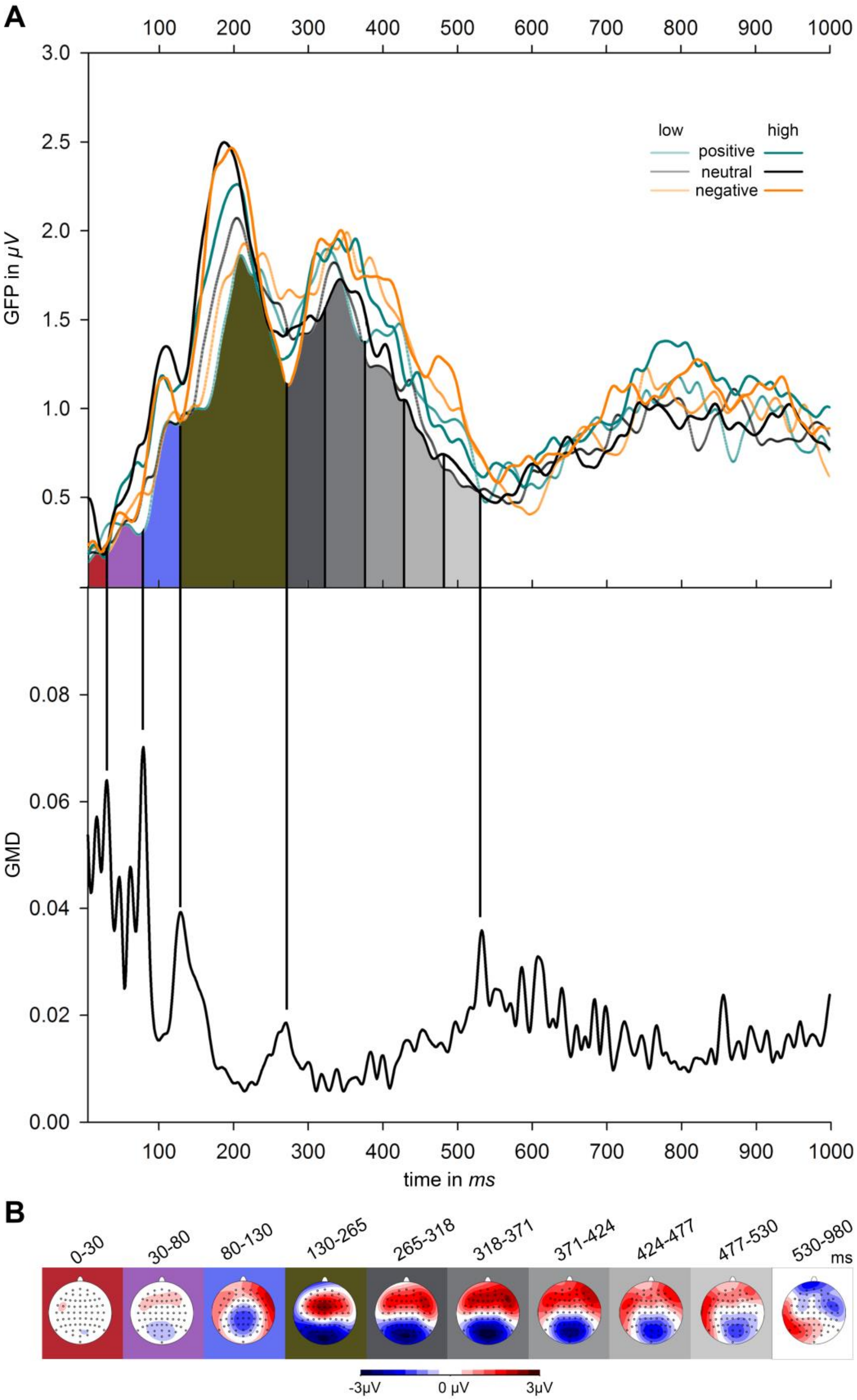

Fig 2.1 | Effects of emotional valence and volume level on electrophysiological parameters. (A) The upper graph shows global field power (GFP) across all participants, contrasted for emotionally positive, negative and neutral written words and high and low level words. The lower graph depicts global map dissimilarity (GMD) averaged across all subjects and experimental conditions. Vertical black lines mark the segment borders, which were defined according to the GMD peaks. Between the peaks at 265 ms and 530 ms, ERPs were divided into five equally long time frames and after the last clear peak at 530 ms in consecutive time frames of 50 ms between up to 980 ms. (B) Maps show the global scalp distribution averaged across all conditions during the time frames flanked by the borders depicted in panel (A).

Results

Performance

Overall, participants performed highly accurate in the one-back task (percent correct = 99.6 %, SD = 1.1).

Effects of volume level

Significant interaction effects of electrode x volume level were revealed in the two consecutive time frames between 80 and 130 ms, $F(63,1764) = 6.314$, $p < .001$, $\eta_p^2 = .184$, and between 130 and 265 ms, $F(63,1764) = 8.948$, $p < .001$, $\eta_p^2 = .242$. These interactions were driven by significant volume level effects in a central ROI (electrodes: C1, C2, Cz, CP1, CP2, CPz, FC1, FC2, FCz). As can be seen in Figure 2.2 (panel C), high volume words elicited more negative amplitudes as compared to low volume words between 80 and 130 ms, $F(1,28) = 45.456$, $p < .001$, $\eta_p^2 = .619$, and more positive amplitudes between 130 and 265 ms, $F(1,28) = 45.453$, $p < .001$, $\eta_p^2 = .614$.

Effects of emotion

The omnibus ANOVA revealed electrode x emotion interaction starting between 371-424 ms, $F(126,3528) = 2.570$, $p < .01$, $\eta_p^2 = .084$, reflecting significant emotion effects in a frontal ROI (electrodes: AF3, AF4, AF7, AF8, AFz, FP1, FP2, FPz), $F(2,56) = 4.623$, $p < .05$, $\eta_p^2 = .142$ and in a parieto-occipital ROI (electrodes: CPz, CP1, CP2, Pz, P1, P2, P3, P4, POz, PO3, PO4, O1, O2, Oz), $F(2,56) = 4.465$, $p < .05$, $\eta_p^2 = .138$. As depicted in panel B of Figure 2.2, spoken words of negative content elicited a stronger relative anterior positivity, $F(1,28) = 13.612$, $p < .01$, $\eta_p^2 = .327$, and parieto-occipital negativity, $F(1,28) = 11.461$, $p < .01$, $\eta_p^2 = .290$, than neutral words, while positive words only showed trends towards significance in the anterior ROI, $F(1,28) = 6.480$, $p = .051$, $\eta_p^2 = .188$.

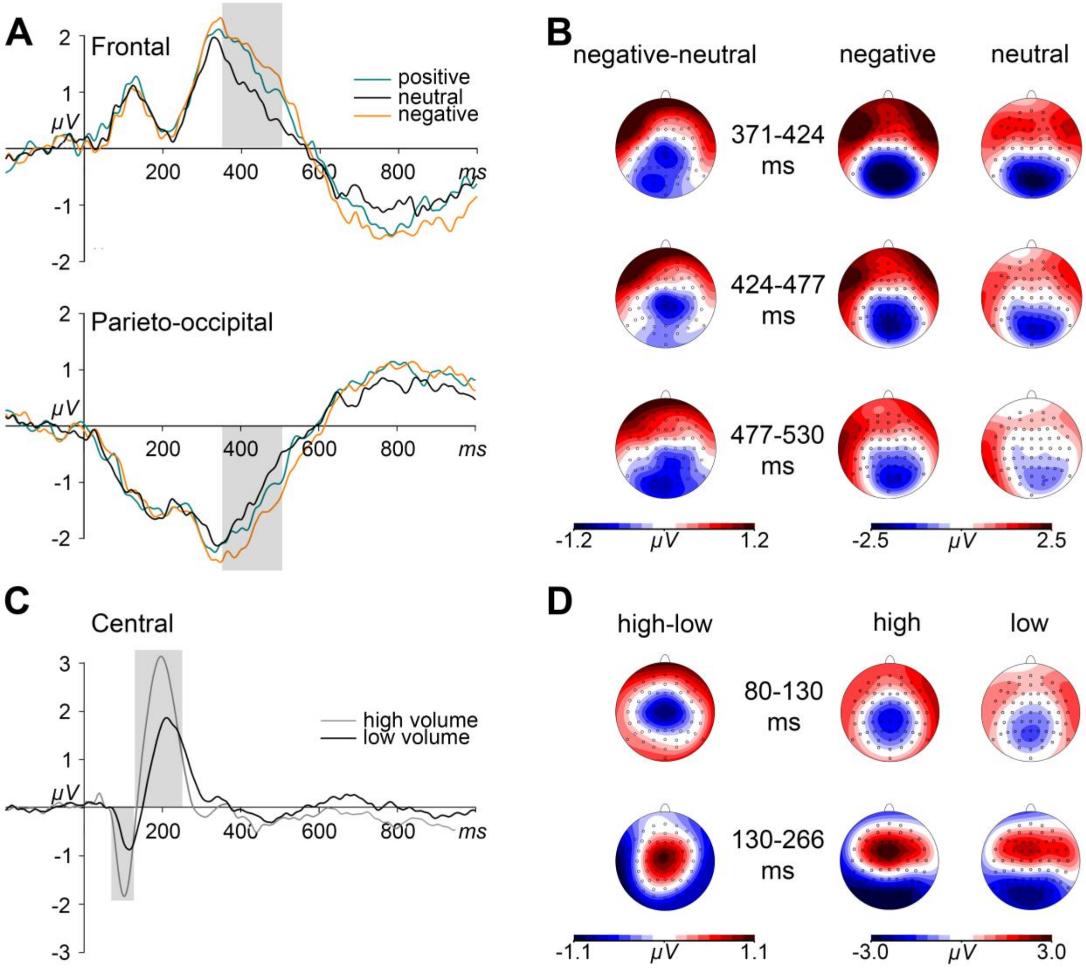

Figure 2.2 | Effects of emotional content and volume level on event-related potentials.
(A) Grand mean ERP waveforms elicited by emotionally positive, negative and neutral words are depicted from frontal and parieto-occipital ROI electrodes. (B) Depicted maps show the scalp distributions of the ERP differences between negative and neutral words within the time intervals of significant emotion effects as well as the distribution of ERPs, separated for negative and neutral words. (C) Grand mean ERP waveforms, contrasted for high and low volume level, are depicted for central ROI electrodes. (D) Maps depict the scalp distributions of ERP differences between high and low volume level words as well as the topographies of ERPs for both volume conditions within the indicated N1 and P2 time intervals.

The emotion x electrode interaction sustained during the two consecutive time frames, i.e. between 424 and 477 ms, $F(126,3528) = 2.046$, $p < .05$, $\eta_p^2 = .068$, and, between 477 and 530 ms, $F(126,3528) = 1.993$, $p < .05$, $\eta_p^2 = .066$. In the first interval, this interaction resulted from significant emotion effects in both the frontal, $F(2,56) = 3.680$, $p < .05$, $\eta_p^2 = .116$, and the parieto-central ROI, $F(2,56) = 3.522$, $p < .05$, $\eta_p^2 = .112$. These effects were driven by larger amplitudes to negative than neutral words in the frontal, $F(1,28) = 7.695$, $p < .05$, $\eta_p^2 = .216$, and the centro-parietal ROI, $F(1,28) = 8.111$, $p < .05$, $\eta_p^2 = .225$. In the following interval (477-530 ms), an emotion effect again was discernible at the frontal as well as at parieto-occipital electrodes, $F(2,56) = 3.406$, $p < .05$, $\eta_p^2 = .108$ and $F(2,56) = 5.761$, $p < .01$, $\eta_p^2 = .171$, respectively. Negative words only showed a trend to elicit a frontal positivity, $F(1,28) = 6.497$, $p = .051$, $\eta_p^2 = .188$ but significant effect at parieto-occipital electrodes, $F(1,28) = 10.975$, $p < .01$, $\eta_p^2 = .282$.

Emotion x electrode interactions did not reach significance in any of the other time frames.

Interaction effects

Importantly, there was no three-way interaction of the factors emotion, volume level, and electrode in any of the time frames. Thus, in the present study, volume level did not modulate emotion effects.

Discussion

The present study aimed at investigating the interplay of volume level and emotional content in spoken words. To this end, words of positive, negative and neutral content in two different volume levels were presented while recording ERPs. As expected, volume level led to a modulation of early processing stages at level of the N1 (80-130 ms) and P2 (130-265 ms) component, confirming the well-known influence of intensity on N1/P2 peak-to-peak amplitude (Rapin et al., 1966; Beagley and Knight, 1967; Adler and Adler, 1991; Thaerig et al., 2008). The processing of words presented at higher volume level led to enhanced N1 and P2 amplitudes at central electrodes compared to low volume level words.

The processing of words of emotional relative to neutral content elicited emotion effects starting around 370 ms after stimulus onset. Negative emotional content led to an increased frontal positivity and a parieto-occipital negativity compared to neutral content between 371 and 530 ms. Only between 371 and 424 ms, a trend was discernible for positive words eliciting an enhanced anterior positivity compared to neutral words. The scalp distribution of these effects showed similarity to an EPN topography to some extent. They might indeed resemble an auditory EPN as it was proposed to exist as an equivalent to the visual EPN (Mittermeier et al., 2011; Jaspers-Fayer et al., 2012). In the visual domain, the EPN component is assumed to reflect a boost in visual encoding due to enhanced attention allocation to emotional stimuli, mainly based on its temporal and topographical similarities to the so-called selection negativity (SN) triggered by voluntary attention allocation (Hillyard and Anllo-Vento, 1998; Schupp et al., 2007).

However, compared to previous, particularly visual EPN effects, the distribution of the emotion effect in the present study expanded further towards central scalp areas. Thus, the effect of emotional content shows more similarity to the N400 component. The N400 is known as an indicator of semantic processing and was reported not only for visual, but also for auditory paradigms (Hahne and Friederici, 2002; Wambacq and Jerger, 2004; Diamond and Zhang, 2016). In the visual modality, it is modulated by the overall expectancy and congruity of (neutral) stimuli in semantic contexts (cf. Kutas and Federmeier, 2011), but also for words of negative content when embedded into sentences (e.g., Holt et al., 2009; Bayer et al., 2010). As in general N400 amplitudes are increased for unexpected compared to expected stimuli, it is reasonable to assume that words of negative content were less expected in the present study. The one-back task employed in the present study required a cross-modal comparison between visually presented catch stimuli and preceding spoken target words. Although the one-back trials incidentally occurred in about 10 percent of all trials, this paradigm might have spanned a very general semantic context, in which spoken words of negative content might have been less expected than words of neutral or positive emotional valence. However, we would like to point out that N400 effects recorded with comparable setups (mainly, average reference) usually occur with a more central maximum. Therefore, the ERP emotion effect in this study might reflect a mixture of both components rather than a solely EPN-like or N400 component.

Interestingly, the factor volume level did not interact with the factor emotion. Although we revealed reliable effects of emotional content, showing anterior positivities and parietal to posterior negativities for emotional compared to neutral words, these effects were not modulated by the loudness of the presented words. This finding might indicate that the mechanism underlying interactions of emotional content and stimulus-triggered attention is acting across different stimulus domains in the visual modality (i.e., both for pictures and written words), but presumably not across different modalities. In the visual domain, comparable interaction effects of stimulus size and emotional content were found on the EPN component in response to emotional pictures and words (De Cesarei and Codispoti, 2006; Bayer et al., 2012a). For the auditory modality, this study was not revealing such an interaction on a component, which might be interpreted as a functional equivalent to the visual EPN. Alternatively, one could assume that actual differences between volume levels might have been too small, since the volume bandwidth was limited to volume levels that were audible, but not too loud, in order to prevent participants from startling. However, participants' reports indicated that volume levels were clearly distinguishable. Modulations of N1 and P2 amplitudes provide further proof that the volume level manipulation in itself was successful.

Importantly, emotion in the auditory modality is conveyed via two different channels: the content of an utterance and the tone of voice, more specifically the prosody, which both impact spoken language processing. Thus, it seems likely that emotional relevance in the spoken modality is not only conveyed by the content but also by the prosody of the utterance (Steinhauer et al., 1999; Wambacq and Jerger, 2004; Kotz and Paulmann, 2007). The words used in the present study were spoken in a neutral tone of voice in order to make them directly comparable to written words, which do not have this second communication channel. Presumably, what makes a heated argument even more emotional and relevant to us is not just a raising of the voice per se, but raising the voice with a meaningful prosody. Raising the voice per se would lead to a change in different acoustic parameters as rhythm, timbre, and pitch, contrasted with a distance-related increase in volume which has been proposed already by gestalt psychologists (e.g., Metzger, 1942). Therefore, manipulating merely the loudness of the neutral spoken stimuli might not have heightened their relevance and interaction of loudness and emotional content might depend on corresponding changes of prosody.

In addition, the social context may play an important role in this paradigm. In the present study, the words were played back via loud speakers without a speaker being visible. However, a recent study by Rohr and Abdel Rahman (Rohr and Abdel Rahman, 2015), provided evidence that a more naturalistic context provided by the presence of a speaker's face on the screen can strongly enhance emotion-related ERPs. Additionally, for visually presented words, previous studies demonstrated that context, especially self-reference and self-other discrimination enhances effects of emotion (Herbert et al., 2011a; Herbert et al., 2011b; Fields and Kuperberg, 2012). Due to these findings, it might be conceivable that interactions between volume level and emotional content might have occurred if the word's relevance had been augmented by self-relevance or by multimodal presentation. A further potential explanation for the absence of interactions between emotion and loudness in the present study is the already mentioned incremental nature of auditory stimuli resulting in poorer synchronization of the EEG signals across stimuli.

The present results hint to a broader difference between visual and auditory language processing than only the impact of volume on emotional processing. Notably, next to the absence of interactions between volume level and emotional content, there were no later main effects of volume level on ERPs. For visual stimuli of different domains, variation of stimulus size was shown to cause the EPN and the LPC; (De Cesarei and Codispoti, 2006; Bayer et al., 2012a). In the present study, no effects of variation in volume level occurred after the early P1-N2 complex. To best of my knowledge, most studies investigated the loudness dependence of the auditory evoked potential (LDAEP) only for the N1-P2 complex (Carrillo-de-la-Peña, 1999; Schadow et al., 2007; Park et al., 2010), but not for later subsequent components. Obviously, volume level impacts early perceptual processing (as reflected in N1 and P2), but might not be considered during high-order processing and stimulus evaluation. Probably, effects of stimulus size on later ERP components in the visual domain are resulting from the closer proximity perception. Presumably, in the auditory modality effects of volume level on higher-order processing components are missing since the proximity manipulation might be depending on more than one factor and, possibly, on visual input as well. Thus, volume level differences which are unaccompanied by an "approaching" visual input might not be sufficient for a proximity manipulation and might therefore not impact high-order processing stages. It is conceivable that the proximity manipulation would have been enhanced by a speaker being present, at different distances for instance. This question should be elaborately addressed in future research. Not only the lack of LPC modulations by volume level, but also the general absence of later emotion effects after the EPN/N400-effect was conspicuous.

In conclusion, the present study revealed effects of emotional content between ~ 370 and 530 ms. The scalp distribution of these emotion effects showed similarity to topographies of an EPN as well as N400. As opposed to prior expectations, no clear statement about the existence of an auditory EPN is actually possible. Furthermore, no emotion effects on later ERP components, as for example the LPC, were shown. To answer the question, whether EPN and LPC have their counterparts in the auditory modality, further studies directly comparing emotion effects in both modalities are required. Importantly, neither the reported emotion effect nor any other investigated time frame showed an interplay of volume level and emotional content. This could indicate that the mechanisms responsible for interactions of emotional content and stimulus-triggered attention in language processing might be limited to the visual modality. However, as discussed above it is conceivable that the non-social, artificial delivery of spoken words in the current experimental setting does not correspond to the more naturalistic context, where the interactions of emotional content and stimulus-triggered attention might be taking place.. Achieving an auditory proximity manipulation in an experimental setup where participants actually see the source of auditory stimuli might be a limitation of the present study. Future research is needed to prove the absence of the interaction effects with further evidence, in particular for different types of emotional auditory stimuli and presentations of speakers at varying distances.

3 Effects of Emotional Content in Written and Spoken Word Processing: Evidence from Event-Related Brain Potentials (Study 2)

Introduction

The use of language is essential to nearly every aspect and interaction in our everyday lives. Language is a complex system that is used to inform the people around us of our inner state, thoughts and feelings, to control another person or event, and to question and understand the world around us. When talking about language, we could refer to speech (spoken language), as well as writing (written language). Both modalities of language differ from each other in several aspects. Speech is usually a dynamic interaction between two or more people, while writing normally does not receive an immediate reply. Written language tends to be more carefully planned and complex (Hammond, 1990), and feedback usually takes more time. While written language is based on letters, which constitute a parallel input with a clear delineation to the reader, speech is based on phonemes constituting a serial input without a clear delineation. The different inputs of spoken and written language are leading to a sensory-specific early processing as shown for single words with fMRI and EEG (Marinkovic et al., 2003). Later on, Marinkovic and colleagues report a multimodal convergence of information from sensory to associative areas, letting us assume a processing of semantic meaning in comparable brain areas.

However, the question remains if emotional semantics interact with this process at the same stages or whether there exists a different system for the detection of emotion during written and spoken word processing. A vast body of research using ERPs investigated the boundary conditions of emotion effects on written word processing, while the spoken modality was neglected. The majority of studies investigating emotion in spoken utterances rather focused on the impact of prosody alone or the interrelation of prosody and semantic information (e.g. Wambacq and Jerger, 2004; Kotz and Paulmann, 2007; Agrawal et al., 2012; Schirmer et al., 2013).

By now, there is some evidence for the existence of auditory counterparts to visual, emotion-related ERPs, such as the EPN and the LPC. An auditory EPN was reported by two studies coming from the same laboratory and specifically investigating the existence of an auditory EPN (Mittermeier et al., 2011; Jaspers-Fayer et al., 2012). Emotion effects on "auditory EPN" components were shown for emotional compared to non-emotional paradigms. However, both stduies contrasted the processing of emotional words and affectively uttered syllables to the processing of neutral tones which is presumably activating different neural processes. Furthermore, the arguments for defining this component as an equivalent to the visual EPN remains surprisingly speculative. Last but not least, in the visual modality EPN effects are post-lexical and some of the reported effects (i.e. around 130 to 150 ms and 250 to 390 ms) are rather early for lexical access to have taken place.

In Study 1, negative compared to neutral meaning elicited effects showing similarity to the visual EPN at scalp topography level of difference waves. However, the emotion effects showed apparent differences to the visual modality. First, compared to typical EPN effects, the distribution of the negativity expanded further towards central scalp areas, therefore showing some similarity with a N400 distribution as well. Second, different latencies were revealed compared to the visual modality, which, however, are explainable by the incremental nature of auditory stimuli and the serial information input compared to a parallel visual input. Lastly, these emotion effects were not interacting with volume level, in contrast to the visual modality, where emotion effects on the EPN were shown to be modulated by stimulus size Thus, it seems questionable whether these effects actually represent an auditory EPN.

Next to the vague claim of an auditory EPN, there is the additional evidence for the existence of a later equivalent, namely the visual LPC. Emotion effects on an auditory LPC were shown not only for spoken utterances with emotional prosody (Paulmann et al., 2013; Schirmer et al., 2013) but also to words with emotional connotation (Ofek et al., 2013; Hatzidaki et al., 2015). In Study 1, no modulation of LPC by emotional meaning of the spoken words was shown. In my opinion, both, the inconclusive results concerning equivalents of visual emotion-related ERP effects in the literature and the results of Study 1 showing little similarity to the visual modality indicate the necessity of a direct comparison of auditory and visual emotional content processing.

To my knowledge, only one study investigated the modulation of event-related potentials occurring to the same written and spoken words by emotional content (Rohr and Abdel Rahman, 2015). This was actually realized in a pre-experiment in order to establish a baseline of effects of emotional meaning for their main experiment, especially due to the little amount of evidence on electrophysiological correlates of auditory emotional words. For the visual presentation, they revealed effects on the EPN, as expected. However, LPC effects were absent for written words. For the auditory presentation of words, no early brain responses were found that may be viewed as auditory counterpart to the visual EPN. In general, emotion effects for auditory words were weaker and mainly characterized by a focal N400-like central negativity.

In my opinion, it is still debatable whether an auditory EPN does actually exist and what the boundary conditions of this component are. The same is applicable for the auditory LPC which was entirely absent in Study 1. The second study of this dissertation is therefore aiming at answering the question of the existence of these counterparts to visual emotion effects by directly comparing the effects of emotional content in written and spoken words. This direct comparison of emotion effects in both modalities using the same stimulus material would allow the dissociation of domain-specific and domain-general processing stages.

The present study consisted of two experiments investigating effects of emotional content during reading of and listening to words of positive, neutral and negative emotional meaning. Both experiments comprised the same word material across the two modalities but differed in the type of voice producing the spoken word material. By means of ERPs, the present study aimed at identifying similarities and differences between emotion effects across both modalities regarding their functional loci. In Experiment 1, spoken words were produced by a computer-generated voice in order to exclude any potential influence of a human (presumably meaning-consistent) prosody, thus ascribing effects solely to emotional meaning. In Experiment 2, the same words were spoken by a female human speaker and thus assumed to have an enhanced social significance to the listener. The same word stimuli were used as in Study 1, which have previously been shown to elicit typical emotion-related ERP components as EPN and LPC in the visual domain (Bayer et al., 2012a). Therefore, we expected to replicate effects of emotional content on the EPN and LPC in the visual modality in both experiments. Due to the rather heterogeneous evidence from previous research on effects of emotional content during auditory word processing, no specific a-priori hypothesis could be derived regarding early ERP components. For elaborate, sustained processing of emotional stimuli at later stages, enhanced LPC amplitudes were expected for words of emotional content not only in the visual but also in the auditory domain (Mittermeier et al., 2011; Ofek et al., 2013; Hatzidaki et al., 2015).

At very early, modality-specific stages of sensory processing, usually indicated by the N1-P2 complex, meaning-based effects of emotion are rather unlikely to occur considering the incremental nature of spoken word processing. Instead, effects of emotional meaning were expected to be reflected in an auditory evoked ERP component equivalent to the visual EPN. If it is assumed that an auditory EPN reflects activation in auditory areas comparable to the boost of visual cortex activity reflected in the visual counterpart (e.g. Schacht and Sommer, 2009a, 2009b; Palazova et al., 2011; Palazova et al., 2013), it appears unlikely to find exactly the same component elicited by emotional meaning during spoken word processing. Hence, it would be conceivable to observe an ERP component reflecting an augmented activation of auditory areas and thus representing a modality-specific equivalent to the visual EPN. However, the results of Study 1 give at least some indication for a similar scalp distribution for difference waves of emotional and neutral stimuli. This could argue for a spreading of activation from auditory areas by spoken words to the visual cortex and the enhancement of activation by emotional content typically related to the visual EPN. It has been shown that the visual cortex is not specifically a unimodal system but shows cross-modal activation by (amongst other non-visual stimuli) auditory stimuli (e.g. Zimmer et al., 2004; Poirier et al., 2005; McDonald et al., 2013; Qin and Yu, 2013). To get a better understanding of the neural foundation of the emotion effects in the auditory modality, ERPs in the present study were recorded with a higher number of electrodes in order to be able to analyze neural sources and therefore make a more profound statement about the effects. Sources were estimated for the merged data of both experiments, separately for the two modalities following a two-step approach. In the first, descriptive approach sources were estimated for the single emotion categories, whereas the second, statistical approach had the purpose to estimate the sources of the difference waves between emotional and neutral content.

Experiment 1

Methods

Participants.

In Experiment 1, data were collected from twenty-five native German speakers (18 women, 7 men; mean age = 22.4 years, SD = 3.4), who had normal or corrected-to-normal vision, and no neurological or psychiatric disorders according to self-report. Twenty-four participants were right-handed; one was left-handed (according to Oldfield 1971). Participation was reimbursed with course credit or 8 euros per hour.

Stimuli.

The seventy-two German nouns used in this study were taken from Bayer and colleagues (2012a). The nouns were previously selected from the Berlin Affective Word List Reloaded (Võ et al., 2009), consisting of three different categories: high-arousing positive, low-arousing neutral, and high-arousing negative words (see Table 3.1 for descriptive statistics). The three categories differed significantly in their valence ratings,

$F(2, 69) = 1362.7$, $p \leq .001$, with lower ratings for negative compared to neutral, $F(1, 46) = 725.7$, $p \leq .001$, and higher ratings for positive compared to both negative, $F(1, 46) = 2446.8$, $p \leq .001$, and neutral words, $F(1, 46) = 727.7$, $p \leq .001$. Positive and negative stimuli were matched for arousal, $F(1, 46) = 1.68$, $p = .202$, and differed both significantly from neutral words, $Fs(1, 46) > 99.0$, $ps \leq .001$. In addition, stimulus categories were controlled for word frequency, number of letters and syllables, and imageability ratings, all $Fs(2, 69) \leq 1$.

The spoken words were generated by synthesized speech production software (Linguatec Voice Reader, Munich, Germany). A female voice was chosen from the software to produce the words. Relevant acoustical parameters were extracted using Praat (Boersma und Weenik 2009) and amplitude was measured with a sound level meter. Stimuli were controlled across categories for mean F0, F0 range, duration, and mean amplitude, all $Fs(2, 69) < 2.5$, $p > .05$ (see Table 3.1 for descriptive statistics).

Table 3.1

Descriptive statistics (Means and Standard Deviations) of visual and auditory stimulus word parameters.

Stimuli		**Positive**		**Neutral**		**Negative**	
		Exp 1	Exp 2	Exp 1	Exp 2	Exp 1	Exp 2
Visual	**Valence**	2.1 (0.2)	2.1 (0.2)	0.3 (0.2)	0.3 (0.2)	-2.0 (0.3)	-2.0 (0.3)
	Arousal	3.3 (0.7)	3.3 (0.7)	1.9 (0.2)	1.9 (0.2)	3.5 (0.5)	3.5 (0.5)
	Imageability	5.4 (1.0)	5.4 (0.8)	5.6 (0.4)	5.6 (0.4)	5.5 (0.6)	5.5 (0.6)
	Letters	6.3 (2.0)	6.3 (1.9)	6.3 (1.2)	6.3 (1.2)	6.4 (2.1)	6.4 (2.1)
	Syllables	2.0 (0.9)	2.0 (0.8)	2.0 (0.8)	2.0 (0.8)	2.1 (1.0)	2.1 (1.0)
	Frequency	27.7 (32.0)	27.7 (32.0)	24.6 (29.2)	24.6 (29.2)	24.8 (20.5)	24.8 (20.5)
Auditory	**Mean F0**	155.0 (24.0)	207.2 (8.8)	157.5 (24.0)	202.4 (6.8)	165.3 (20.8)	205.3 (8.5)
	F0 Range	64.4 (33.5)	61.4 (24.2)	78.3 (31.2)	66.0 (14.1)	93.8 (82.8)	57.3 (14.9)
	Duration	656.6 (150.2)	682.2 (123.6)	592.3 (129.2)	628.5 (99.3)	672.4 (167.4)	694.6 (149.3)
	Mean Amplitude	62.3 (3.0)	59.5 (2.5)	64.2 (2.6)	59.4 (2.4)	63.7 (3.3)	59.4 (3.3)

Notes. For ratings, the ranges are: -3 to +3 (valence), 1–5 (arousal), 1–7 (imageability). Frequency is indicated as occurrence per 1 million words; Mean F0, F0 range and duration were measured in Praat and are indicated in Hertz and milliseconds, respectively. Mean amplitude is given in Dezibel.

Procedure.

The study was approved by the ethics committee of the Department of Psychology at the University of Goettingen, Germany, and was conducted according to the Declaration of Helsinki. Before the beginning of the experiment, participants were acquainted with the experimental procedure and signed an informed consent. After preparation for the EEG recording, participants were seated in a dimly lit, sound-attenuated chamber, facing a computer monitor at a distance of approximately 100 cm. All word stimuli were presented both in the visual and acoustic modality. In the visual modality, the words were presented in the center of a screen in black letters (Arial font, 46-point) on a light gray background. Vertical size of the letters was at most 0.63° and horizontal size of 6.11° of visual angle. In the acoustic modality, the words were presented via loudspeakers (mean loudness = 63.4 dB; SD = 3.02).

Visual and auditory presentations were realized in two separate blocks in counterbalanced order. Each block contained two differently randomized sequences of all 72 trials, resulting in 144 trials per block. All trials had a length of 4000 ms. Visual trials started with a fixation cross, followed by the written word, both presented for 1000 ms. After the word disappeared, a blank screen was presented for the remaining 2000 ms. In acoustic trials, a fixation cross was presented for 3000 ms and the spoken word was played after 1000 ms. After the fixation cross disappeared, a blank screen was presented for 1000 ms. The fixation cross was presented before and during complete stimulus presentation in order to avoid eye movements and offset potentials from the disappearing visual stimulus while the acoustic stimulus was still playing.

Participants were instructed to attentively read or listen to the words, respectively. In order to ensure that subjects were paying attention, a 1-back task was randomly interspersed after 2 to 15 trials by showing a green frame with a question mark on the screen. Subsequently, a word was presented, and participants had to decide by button press whether the presented word was identical or not to the preceding word. These 1-back items were matched in modality to the one of the words in the given block (auditory or visual). After every

27th trial (approximately every 2 minutes) a break was interspersed, which was terminated by the participants via button press.

EEG recording.

The electroencephalogram was recorded from 128 electrodes, mounted in an electrode cap (Easy-Cap, Biosemi, Amsterdam, Netherlands). The common mode sense (CMS) active and the driven right leg (DRL) passive electrode were used as reference and ground electrodes, respectively (cf. www.biosemi.com/faq/cms&drl.htm). Four external electrodes were placed laterally and inferior to the eyes to record eye movements and blinks. Electrode offsets were kept below a threshold of +/-20mV. Signals were recorded at a sampling rate of 512 Hz and a bandwidth of 104 Hz.

Data analysis.

Offline, data was processed with the BrainVision Analyzer (Brain Products GmbH, Munich, Germany). EEG signals were average-referenced and band-pass filtered (Butterworth 0 phase filter) between 0.032 (time constant 5 s, 12 dB/octave) and 40 Hz (48 dB/octave). A notch-filter (50 Hz) was additionally applied. Data were corrected for blinks and eye movements using Surrogate Multiple Source Eye Correction (MSEC; Ille et al., 2002) as implemented in BESA (Brain Electric Source Analysis, MEGIS Software GmbH, Gräfeling, Germany). The continuous EEG signal was segmented into epochs of 1200 ms starting 200 ms before stimulus onset and referred to a 100 ms pre-stimulus baseline. For some participants, electrodes with poor signal (1.19 %) were interpolated by spherical splines in BrainVision Analyzer (Order of splines: 4; Maximal degree of Legendre Polynomials: 10, Lambda: 1E-05). After discarding epochs containing artifacts (voltage steps larger than 50 µV, 200µV/200ms intervals difference of values, amplitudes exceeding -150 µV/150 µV, and activity smaller than 0.5 µV), ERP segments were averaged per participant and experimental condition (visual/acoustic, positive/neutral/negative, resulting in 3 experimental conditions per modality).

The measurement of ERP amplitudes was performed in time frames segmented according to visual inspection of measures of global field power (Lehmann and Skrandies, 1980) and global map dissimilarity (Brandeis et al., 1992), separately for the two modalities. Figure 3.1 depicts GFP contrasted for emotion conditions (positive, neutral, negative) and GMD, which has been calculated across the three conditions. GFP reflects the overall ERP activity across the scalp at any given moment. GMD reflects the dissimilarity between scalp topographies of adjacent time points and demarcates the borders between periods of relatively stable topographies indicating continued processing within similar brain areas. These transition times were used as the limits of the time segments for which mean ERP amplitudes were calculated. GMD peaks were clearly observable at the following time points for written words at 0, 70, 120, 180, 234, and 350 ms, and for spoken words at 0, 75, 138, 250, 303, and 515 ms. As can be seen in the bottom panel of Figure 3.1, these segments converge with distinguishable scalp field maps over time. In order to allow for fine-grained analyses of ERPs after the last clear segment borders indicated by GMD, mean ERPs were analyzed in consecutive time frames of 50 ms between 350 and 800 ms in the visual modality and between 515 and 815 ms in the auditory modality, respectively. For the same reason, auditory ERPs between 303 and 565 ms were additionally sub-segmented into five time intervals of equal length (53 ms each; please see Figure 3.1 for details).

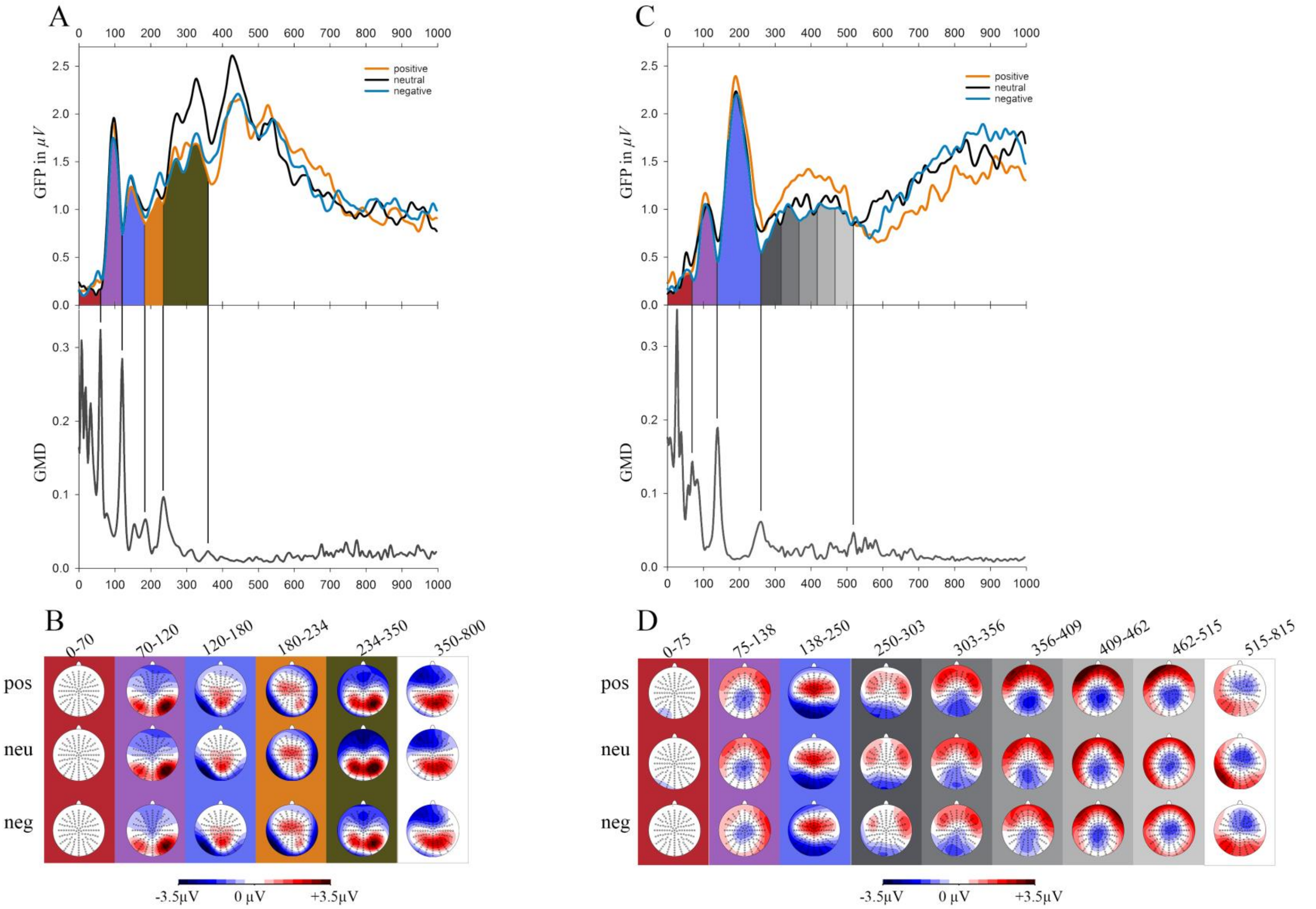

Figure 3.1 | Effects of emotional valence on electrophysiological parameters for written (left panel) and spoken words (right panel) in Experiment 1. **(A)** The upper graph shows global field power (GFP) across all participants, contrasted for emotionally positive, negative and neutral written words. The lower graph depicts global map dissimilarity (GMD) across all subjects and averaged across positive, negative and neutral written words. Vertical black lines mark the segment borders, which were defined according to the GMD peaks. **(B)** Maps show the scalp distribution separately for positive (pos), neutral (neu) and negative (neg) written words during the time windows flanked by the borders depicted in **(A)**. **(C)** GFP and GMD plots for spoken words. Between the peaks at 250 ms and 515 ms, ERPs were divided into 5 equally long time windows (green). **(D)** Maps show the scalp distribution separately for positive (pos), neutral (neu) and negative (neg) spoken words during the time windows flanked by the borders depicted in **(C)**.

Mean ERP amplitudes within these time borders were averaged across electrodes (N = 7) in 16 separate clusters. Amplitude differences were assessed by repeated-measures ANOVAs, separate for the two modalities, including the within subject factors emotion (3 – positive, neutral, negative) and cluster (16). Degrees of freedom in ANOVAs were adjusted using Huynh–Feldt corrections. If indicated by significant emotion by cluster interactions in these exploratory analyses, effects of emotion were further tested in regions of interest (ROIs) that were defined based on visual inspection of the ERP difference waves within the specific time frames. For post-hoc comparisons, *p*-values were Bonferroni adjusted.

Results

Performance

Overall, participants showed very high accuracy in the 1-back task (percent correct = 99.38 %, SD = 0.01).

Visually evoked potentials

In the visual domain, a significant interaction effect of cluster x emotion was revealed in a typical EPN time frame between 234-350 ms $F(30,720) = 4.493$, $p \leq .001$, $\eta_p^2 = .158$, driven by significant emotion effects in a posterior ROI (electrodes: A9, A10, A11, A12, A13, A14, A15, A23, A24, A25, A26, A27, A28, B6, B7, B8, B9, B10, B11, B12, B14, B15, D24, D25, D30, D31, D32), $F(2, 48) = 8.872$, $p \leq .001$, $\eta_p^2 = .270$, as well as in an anterior ROI (electrodes: C2, C11, C12, C13, C14, C15, C18, C19, C20, C21, C22, C23, C24, C25, C26, C27, C28, D2), $F(2, 48) = 5.611$, $p < .05$, $\eta_p^2 = .189$. As depicted in Figure 3.2 (panels A and B), both positive and negative words elicited stronger relative posterior negativities, $F(1, 24) = 14.862$, $p < .01$, $\eta_p^2 = .382$ and $F(1, 24) = 22.916$, $p < .001$, $\eta_p^2 = .488$, and anterior positivities compared to neutral words, $F(1, 24) = 6.687$, $p < .05$, $\eta_p^2 = .218$ and $F(1, 24) = 12.351$, $p < .01$, $\eta_p^2 = .340$.

The emotion by cluster interaction sustained during the two immediately following time frames, i.e. between 350 and 450 ms, $F(30, 720) = 3.350$, $p < .01$, $\eta_p{}^2 = .122$ and $F(30,720) = 2.483$, $p < .05$, $\eta_p{}^2 = .094$. As can be seen in Figure 3.2, these emotion effects showed high similarity to the EPN effect in the preceding interval. This impression was verified by significant emotion effects in the posterior ROI, $F(2, 48) = 4.846$, $p < .05$, $\eta_p{}^2 = .168$ and $F(2, 48) = 5.213$, $p \leq .01$, $\eta_p{}^2 = .178$. Pairwise comparisons revealed significant ERP amplitude differences consisting of enhanced amplitudes for positive compared to neutral words, $F(1, 24) = 11.811$, $p < .01$, $\eta_p{}^2 = .330$ and $F(1, 24) = 14.486$, $p < .01$, $\eta_p{}^2 = .376$. Emotion effects in the anterior ROI were less pronounced within these time frames, $F(2, 48) = 2.734$, $p = .078$, $\eta_p{}^2 = .102$, and, $F(2, 48) = 2.817$, $p = .070$, $\eta_p{}^2 = .105$.

The emotion by cluster interaction failed significance in any of the other time frames, i.e. prior to the EPN as well as during subsequent intervals that would resemble classical LPC time frames (450-600 ms).

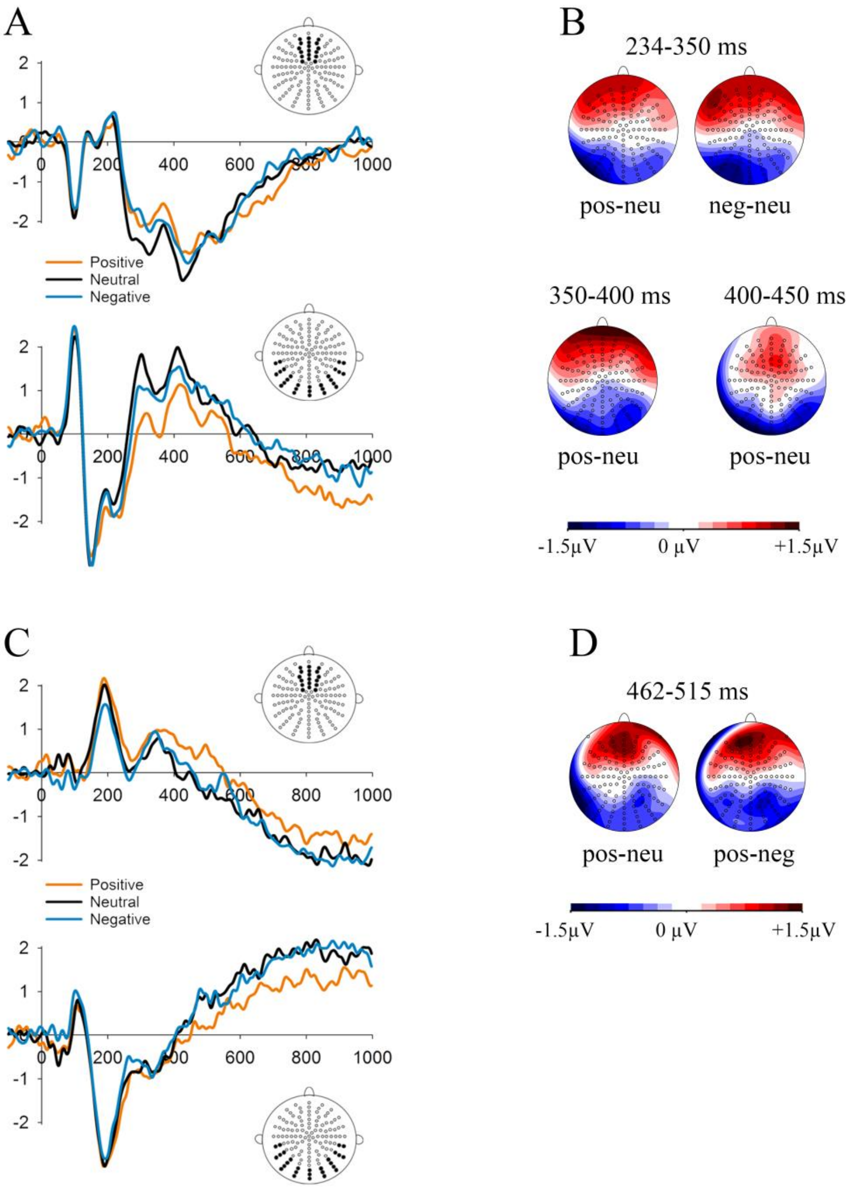

Figure 3.2 | Effects of emotion on ERPs for written words (upper panel) and spoken words (lower panel) in Experiment 1. **(A)** Grand mean ERP waveforms elicited by emotionally positive, negative and neutral written words are depicted as the averaged activity in an anterior and a posterior region of interest. **(B)** Scalp distributions of significant emotion effects are depicted as differences between indicated emotion categories in depicted time frames. **(C)** Grand mean ERP waveforms to positive, negative and neutral spoken words. **(D)** Difference maps for emotion categories of spoken words showing significant emotion effects.

Auditory evoked potentials

For the acoustic modality, the omnibus ANOVA revealed a significant emotion by cluster interaction between 462 and 515 ms, $F(30, 720) = 2.624$, $p < .05$, $\eta_p^2 = .099$, due to significant emotion effects in the anterior ROI, $F(2, 48) = 6.118$, $p < .01$, $\eta_p^2 = .203$, and the posterior ROI, $F(2, 48) = 3.307$, $p < .05$, $\eta_p^2 = .121$. As depicted in Figure 3.2 (panels C and D), positive words elicited an enhanced anterior positivity as compared to both neutral, $F(1, 24) = 14.877$, $p < .01$, $\eta_p^2 = .383$, and negative words, $F(1, 24) = 7.185$, $p < .05$, $\eta_p^2 = .230$. At posterior electrode sites, the relative negativity for positive words compared to neutral and negative words were discernible only as trends, $F(1, 24) = 5.668$, $p = .078$, $\eta_p^2 = .191$ and $F(1, 24) = 4.859$, $p = .111$, $\eta_p^2 = .168$, respectively.

The emotion by cluster interaction did not reach significance in any of the other investigated time frames.

Discussion

The present experiment directly compared effects of emotional content across two different modalities of word processing. To this aim, participants heard or read the same high-arousing nouns of positive or negative valence as well as low-arousing neutral nouns.

Written nouns of emotional content elicited an enhanced negativity at posterior electrode sites and – and as its counterpart – an anterior positivity between 230 and 350 ms, thus resembling the typical EPN scalp distribution and latency, with residual effects lasting up to 450 ms. These results replicate the main finding reported by Bayer et al. (2012a), who used the same word stimuli and are further in good accordance with previous studies concerning the visual EPN (Kissler et al., 2007; Kissler et al., 2009; Schacht and Sommer, 2009a, 2009b; Bayer and Schacht, 2014). Although the EPN was initially observable for both emotionally positive and negative words, this ERP modulation sustained longer for positive than negative valence. Such a bias for positive valence ('positivity bias') has often been reported being evident in ERPs (Herbert et al., 2006; Kissler et al., 2009; Schacht and Sommer, 2009a; Bayer et al., 2012a; Bayer and Schacht, 2014), but also in behavioural indicators (e.g., Schacht and Sommer, 2009b) and even in amygdala activity (Herbert et al., 2009).

During spoken word processing, ERP effects of emotional content occurred at longer latencies, compared to the visual domain, starting around 450 ms. This delay appears reasonable since auditory information is delivered over time, which is in contrast to the immediate presentation of visual words where the entire perceptual information is presented at once. Interestingly, this ERP effect was mainly driven by emotionally positive words, similar to the 'positivity bias' in visual word processing. Another similarity between the emotion effects is worth mentioning: In both modalities, the distribution of ERP differences between emotional and neutral words consisted of an enhanced negativity at posterior electrodes sites and a corresponding polarity reversal over fronto-central regions. This finding was rather surprising since the topographies of the single emotion categories underlying this ERPs showed a rather different distribution in both modalities. However, it should be noted that even though emotion effects at difference level show a comparable scalp distribution across modalities, the size of the emotion effects at anterior and posterior sites point to differences across modalities: Whereas the processing of emotional content for spoken words led to stronger effects at anterior sites compared to the posterior effect, the effects for written word processing were stronger at posterior electrode sites.

Under the experimental conditions realized here, modulation of the LPC component was absent in both modalities. Previous studies have reliably demonstrated that emotion-related LPC augmentations in the visual domain are highly context- and task-dependent (Fischler and Bradley, 2006; Schacht and Sommer, 2009b, 2009a). Therefore, it seems conceivable that for the 1-back task implemented in the present study, a deep lexico-semantic processing of the words was not strictly task-relevant. For spoken words, Rohr and Abdel Rahman (2015), who implemented a similar task as in the present study, recently reported an absence of LPC modulations by the processing of words' emotional content. Importantly, however, they demonstrated that the mere presence of a speaker's face could massively facilitate early emotion-related ERPs in terms of shorter latencies, augmented amplitudes and longer duration. Considering this, one might assume that the ERP effects of emotional content in the present experiment could be more pronounced if the speaker was of higher social relevance or more naturalistic compared to a computer-generated artificial voice. Therefore, we aimed at replicating and extending the findings of Experiment 1, in particular within the auditory modality, with a follow-up experiment wherein the words were spoken by a professional human female speaker. Hence, we expected similar ERP effects for emotional compared to neutral words in both modalities but of presumably larger amplitudes for spoken words in Experiment 2.

Experiment 2

Methods

Participants

Data were collected from twenty-four participants; two datasets had to be discarded due to excessive ERP artifacts. All remaining participants (N = 22; 14 female; mean age = 21.5 years, SD = 2.5) fulfilled the same criteria as participants of Experiment 1. Nineteen participants were right-handed; three were left-handed (Oldfield, 1971). Participation was reimbursed with course credit or 8 euros per hour.

Stimuli

For the auditory presentation, all words were recorded from a female speaker who was requested to produce all words with emotionally neutral prosody. As in Experiment 1, stimulus categories were controlled for mean fundamental frequency (F0), F0 range, duration and mean amplitude, all $Fs(2, 69) < 2.2$, all $ps > .05$ (see Table 3.1).

Procedure

For Experiment 2, exactly the same procedure and experimental design were used as in the first experiment. Only a small adaption was made concerning the 1-back task. Here, the

1-back task test stimuli – either the same word as presented before or a new distracter word – were always presented visually. Stimuli in the auditory block were played with a mean loudness of 59.4 dB (SD = 2.25).

EEG recording

Data recording followed the same procedure as in Experiment 1.

Data analysis

For segmentation of averaged ERPs into components the same procedure as in Experiment 1 was applied. As can be seen in Figure 3.3, visual inspection of GMD and GFP measures revealed time segment borders for ERPs to written words at 0, 70, 124, 170, 230, and 350 ms and for spoken words at 0, 80, 135, 260, 311, 362, 413, 464, and 515 ms. Highly similar to Experiment 1, no further clear GMD peaks appeared after 350 and 515 ms, respectively. Thus, following our previous analyses, ERPs were quantified in consecutive time segments of 50 ms after these two final GMD-defined borders.

As becomes obvious from a direct comparison of the ERP data depicted in Figures 3.1 and 3.3, ERP component structures over time showed high similarity between the two experiments in terms of dynamics and topographies. Therefore and in order to avoid extensive multiple testing, we directly tested emotion effects in both domains using the same ROIs and similar time frames, i.e. 230-350, 350-400, and 400-450 ms for written words, and 464-515 ms for spoken words.

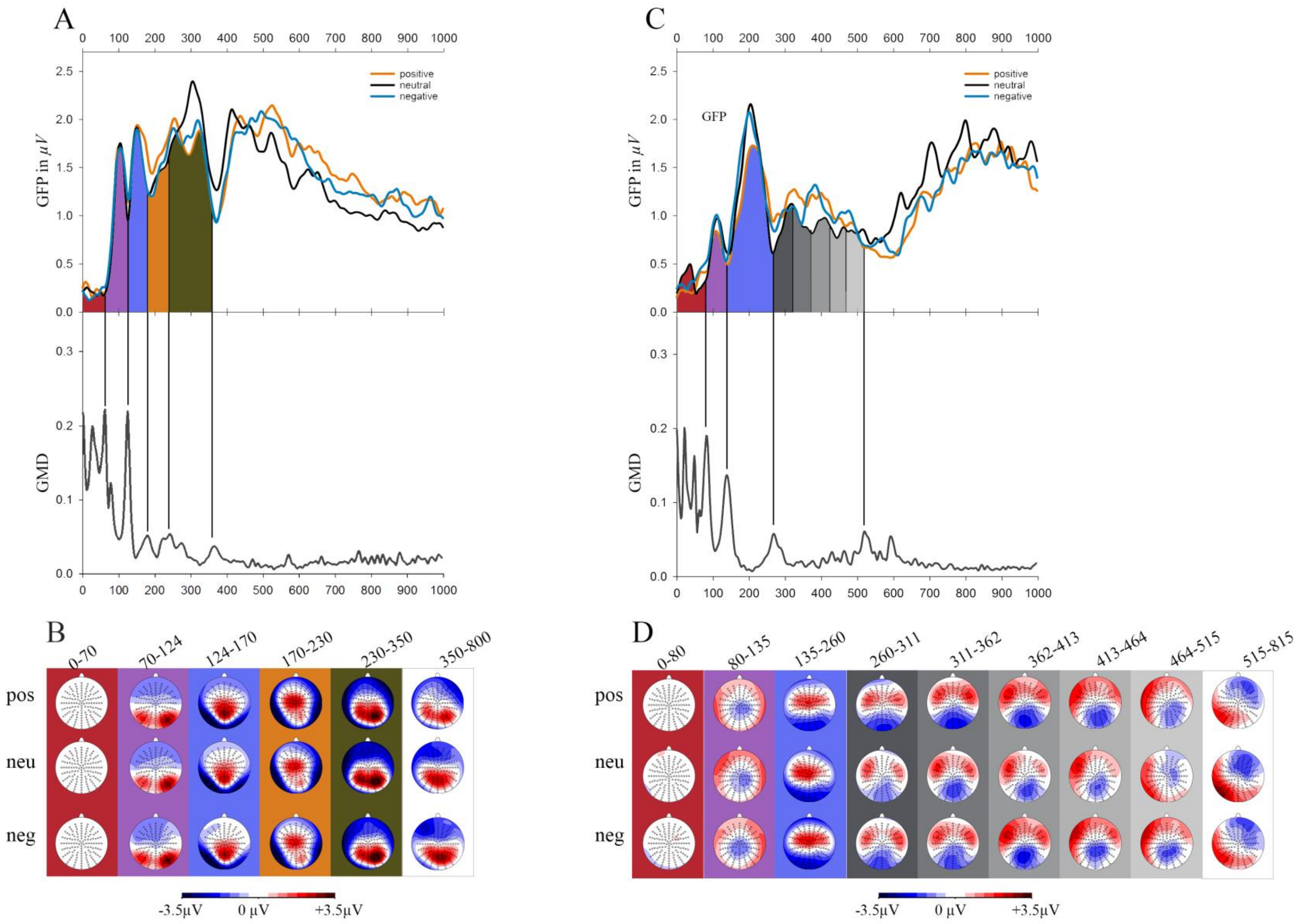

Figure 3.3 | Effects of emotional valence on electrophysiological parameters for written (left panel) and spoken words (right panel) in Experiment 2. **(A)** The upper graph shows global field power (GFP) across all participants, contrasted for emotionally positive, negative and neutral written words. The lower graph depicts global map dissimilarity (GMD) across all subjects and averaged across positive, negative and neutral written words. Vertical black lines mark the segment borders, which were defined according to the GMD peaks. **(B)** Maps show the scalp distribution separately for positive (pos), neutral (neu) and negative (neg) written words during the time frames flanked by the borders depicted in **(A)**. **(C)** shows GFP and GMD plots for spoken words. Between the peaks at 260 ms and 515 ms, time course was divided into 5 equally long time frames (green). **(D)** Maps show the scalp distribution separately for positive (pos), neutral (neu) and negative (neg) spoken words during the time frames flanked by the borders depicted in **(C)**.

Results

1-back Task

As in Experiment 1, participants showed very high accuracy rates (percent correct = 99.43 %, SD = 0.01) in the 1-back task.

Visually evoked potentials

In the EPN time frame 230-350 ms, the ROI analysis revealed emotion effects consisted in an enhanced negativity at posterior electrodes and – as their counterpart – in an enhanced anterior positivity, thus resembling the typical EPN distribution (see Figure 3.4). Emotion effects were significant in both the posterior ROI, $F(2, 42) = 12.296$, $p < .001$, $\eta_p^2 = .369$, and the anterior ROI, $F(2, 42) = 6.384$, $p < .01$, $\eta_p^2 = .233$. Both positive and negative words elicited relative posterior negativities as compared to neutral words, $F(1, 21) = 16.019$, $p < .01$, $\eta_p^2 = .433$ and $F(1, 21) = 10.159$, $p < .05$, $\eta_p^2 = .326$. Furthermore, there was a significant difference between positive and negative words, $F(1, 21) = 7.614$, $p < .05$, $\eta_p^2 = .266$. The anterior emotion effect consisted of larger amplitudes for positive than neutral words, $F(1, 21) = 8.754$, $p < .05$, $\eta_p^2 = .294$.

An emotion effect sustained during the next two time frames (350-400 ms and 400-450 ms) still showing a comparable scalp distribution, however, becoming less distinct. Between 350-400 ms, analyses of the posterior negativity revealed a significant effect of emotion, $F(2,42) = 6.527$, $p < .01$, $\eta_p^2 = .237$, reflecting larger amplitudes for positive compared to neutral words, $F(1,21) = 8.671$, $p < .05$, $\eta_p^2 = .292$. A significant effect of emotion also became evident for the anterior positivity, $F(2,42) = 5.375$, $p < .05$, $\eta_p^2 = .204$; however, post-hoc tests failed significance. Between 400 and 450 ms, the emotion effect was restricted to the posterior ROI, $F(2,42) = 3.521$, $p < .05$, $\eta_p^2 = .144$, however, post-hoc tests failed significance.

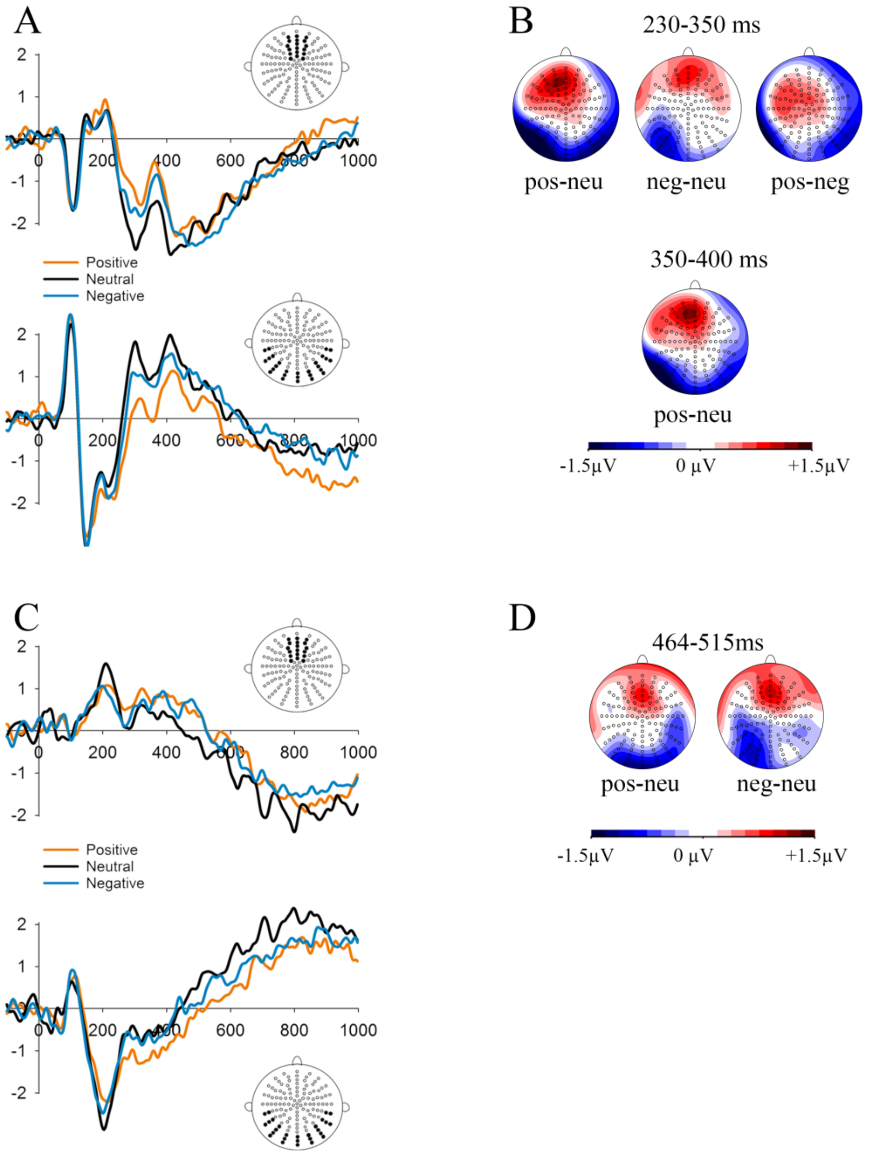

Figure 3.4 Effects of emotion on ERPs for written words (upper panel) and spoken words (lower panel) in Experiment 2. (A) Grand mean ERP waveforms elicited by emotionally positive, negative and neutral written words are depicted from the averaged activity in an anterior and a posterior region of interest. (B) Scalp distributions of significant emotion effects are depicted as differences between indicated emotion categories in depicted time frames. (C) Grand mean ERP waveforms to positive, negative and neutral spoken words. (D) Difference maps for emotion categories of spoken words showing significant emotion effects.

Auditory evoked potentials

In the analyzed time frame between 464-515 ms, emotion effects became evident at the posterior ROI, $F(2, 42) = 4.379$, $p < .05$, $\eta_p^2 = .173$, due to a larger negativity for positive compared to neutral words, $F(1, 21) = 7.503$, $p < .05$, $\eta_p^2 = .263$. Analyses of the anterior ROI revealed an emotion effect, $F(2, 42) = 4.469$, $p < .05$, $\eta_p^2 = .175$. Post-hoc tests indicated these effects consisting of an enhanced positivity to negative compared to neutral words, $F(1, 21) = 7.504$, $p < .05$, $\eta_p^2 = .263$ and a trend for positive compared to neutral words, $F(1, 21) = 6.156$, $p = .066$, $\eta_p^2 = .227$.

Discussion

The aim of Experiment 2 was to replicate the effects for emotional words of Experiment 1, using acoustic word stimuli spoken by a human voice. Moreover, since language spoken by a human voice should be more natural and salient to human listeners, stronger emotion effects were conceivable. For visually presented words, the same emotion-related ERP effects as in Experiment 1 were expected.

Overall, we indeed replicated effects of emotional meaning for written words as reflected in the EPN component at almost the same latency and with similar duration as in Experiment 1 and in a previous study (Bayer et al. 2012a), in which the same word material was used. In accordance with Experiment 1, mainly positive words elicited enhanced relative negativities at posterior and enhanced positivities at anterior electrodes compared to neutral words, while differences between negative and neutral words were much weaker.

For the auditory modality, similar effects were expected as in Experiment 1, however with presumably enhanced amplitudes, since words were produced by a natural human voice and thus of potentially higher biological salience. Indeed, emotion effects for spoken words showed similar scalp topography, consisting of a relative posterior negativity and anterior positivity, as revealed for computer-generated words in Experiment 1. Positive words again elicited enhanced negativities at posterior electrodes and enhanced anterior positivities compared to neutral words. In line with Experiment 1, emotion effects for visual words were more pronounced at posterior electrode sites compared to anterior sites; however, a topographic shift to stronger anterior effects was not as clear as in Experiment 1. The hypothesis of potentially augmented emotion-related ERP effects due to increased salience was therefore not confirmed.

As discussed before, emotion effects in the auditory modality bore a surprisingly high resemblance to the known EPN-effects in the visual modality. However, it remains unclear, whether this resemblance does confirm the existence of an equivalent to the visual EPN and a comparable boost in auditory areas. In order to get a more elaborate understanding of the emotion effects and their similarities across modalities, source localization was performed on the collapsed data of both experiments, which will be described in the next paragraph.

Source localization

To estimate the neural generators underlying the dominant voltage topographies identified at the scalp level, sLORETA (Pascual-Marqui, 2002) was used. sLORETA is a distributed linear inverse solution based on the neurophysiological assumption of coherent co-activation of neighboring cortical areas that are known to have highly synchronized activity. Accordingly, it estimates multiple simultaneously active sources without any a-priori assumption on the number and position of the underlying dipoles (for a mathematical validation of this localization technique, see Sekihara et al., 2005). sLORETA solutions are computed within a three-shell spherical head model co-registered to the MNI152 template (Mazziotta et al., 2001). The source locations are therefore given as (x, y, z) coordinates. sLORETA estimates the 3-dimensional intracerebral current density distribution in 6239 voxels (5 mm resolution). The transformation matrix obtained with this interpolation was applied to the data extracted from Brain Vision Analyzer, converting electric potential differences to standardized current density in the brain (signal-to-noise ratio parameter used in this study = 10).

In order to statistically evaluate the sources that sLORETA calculated for the ERP waveforms, a two-step approach was followed. Given the similarity between the experiments in design, parameters, and ERP results, data from both experiments were merged in order to maximize power and reliability of the source analyses, meaning that a sample of 47 participants was used. The significant activations were identified for emotion categories separately (positive, neutral, negative) within time frames showing significant and stable ERP effects of emotion in both experiments. Therefore, for the visual modality the mean activity was localized in the interval 234-350 ms; and between 463-515 ms for the auditory modality. The sLORETA software package was used to perform the non-parametric statistical analyses. The software estimate, via randomization, the empirical probability distribution for the max-statistic (e.g. the maximum of a *t*-statistic), under the null hypothesis. Due to the non-parametric nature of the method, its validity need not rely on any assumption of Gaussianity (Nichols and Holmes, 2002).

Approach 1

Methods and Results

In the first step, the statistical map of the activation patterns was computed separately for visual and auditory modality, in order to define which brain networks reliably contributed to the ERP components observed on the surface (one-sample T-test against zero for the interval indicated by the scalp analyses, smoothing parameter = 0). The bullet-proof, corrected critical thresholds for the T-test were obtained through 5000 randomizations in SnPM (Statistical non-Parametric Mapping) and then used to define significantly activated clusters. The plots therefore only show voxels that are significantly activated in each condition (positive, neutral and negative words separately), already corrected for multiple comparisons (i.e., for the collection of tests performed for all voxels). Results are summarized in Figure 3.5 and in Table 3.2. As can be seen in Figure 3.5, the visual and acoustic domain differed mainly in the lateralization of the effect. For the visual domain the strongest reliability of activation can be found in the left hemisphere, around the superior parietal lobule (SPL, BA 7, and postcentral gyrus, BA 5), with additional clusters in the inferior parietal lobule (IPL, BA 40), also mostly in the left hemisphere. On the contrary, for the auditory modality the highest T-values and biggest clusters of activation are found in the right hemisphere, around the IPL, with some spreading to the SPL and neighboring areas (for example the cingulated gyrus, BA 31). Moreover, it is noteworthy that in the visual modality all word types elicited activations in the EPN time frame that exceeded the critical thresholds to a larger extent than in the acoustic modality, resembling the scalp ERP data, where the GFP in the EPN frame was noticeably larger for the visual task. Lastly, it can be noted from Figure 3.5 as well as from the Table 3.2 that the reliability of the effects is higher for the emotional categories (both positive and negative) than for the neutral words consistently in both modalities, confirming what the ERP analyses showed at the scalp level.

Table 3.2

Source Localization with sLORETA of both experiments separated for modality and emotion category

	Visual Domain (234-350 ms)				**Acoustic Domain (462-515 ms)**			
	Critical T threshold	**Max T-value**	**Lobe, Area**	**Cluster size (voxels)**	**Critical T threshold**	**Max T-value**	**Lobe, Area**	**Cluster size (voxels)**
Positive	3.25	18.41	Parietal Lobe, BA5	87	3.36	12.37	IPL, BA 40 (RH > LH)	181
		17.14	Bilateral SPL, BA 7	267		9.16	Parietal Lobe, BA5	21
		16.58	Bilateral IPL, BA 40	352		6.65	Right SPL, BA 7	31
		13.58	Bilateral Frontal, BA 6	253				
Neutral	3.22	14.75	Parietal Lobe, BA5 (LH > RH)	87	3.37	5.73	Bilateral Frontal, BA 6	120
		14.63	Bilateral SPL, BA 7	293		5.65	IPL, BA 40 (RH > LH)	93
		13.63	Bilateral IPL, BA 40	368		5.39	Cingulate gyrus, BA 31	105
		12.62	Temporal lobe, Insula, BA 13 (RH > LH)	120		4.44	SPL, BA 7 (RH > LH)	36
Negative	3.27	21.93	Parietal Lobe, BA5 (LH > RH)	90	3.33	7.55	Right IPL, BA 40	117
		21.09	SPL, BA 7 (LH > RH)	286		6.03	Right Parietal Lobe, BA5	22
		19.17	Bilateral IPL, BA 40	357		5.50	Right SPL, BA 7	36
		11.95	Bilateral Frontal, BA 6	219				
		11.36	STG BA 39	107				

Data of Experiment 1 and 2 was merged, in order to maximize power and reliability of the source analyses. The significant activations were identified with non-parametric statistical analyses as implemented in the sLORETA software package. Activation were identified for emotion categories separately (positive, neutral, negative) within time windows showing significant and stable ERP effects of emotion.

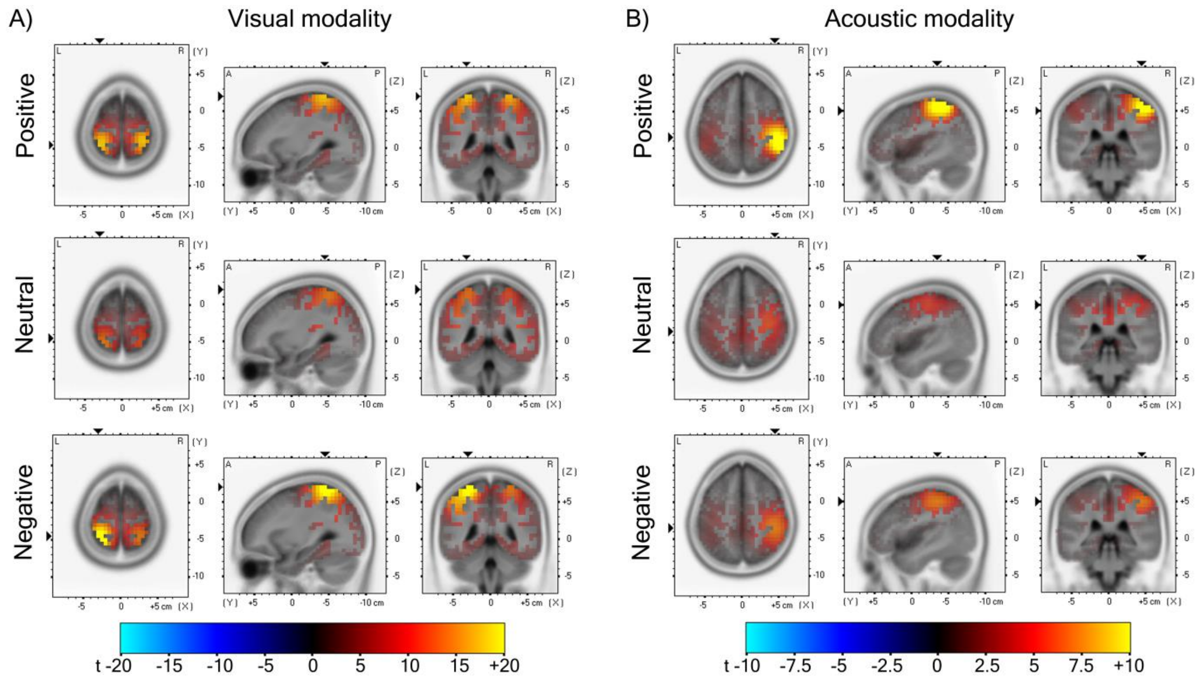

Figure 3.5 Source localization results, based on sLORETA. **(A)** Estimated neural generators during the visual EPN time frame from 230 to 350 ms, separately for the three emotion categories. Highest activation is shown in the left hemisphere, around the superior parietal lobule (SPL, BA 7, and postcentral gyrus, BA 5), with additional clusters in the inferior parietal lobule (IPL, BA 40). **(B)** Estimated neural generators during the auditory EPN time frame from 463 to 515 ms. Clusters of activation are shown in the right hemisphere, around the IPL, the SPL and neighboring areas (e.g. cingulate gyrus, BA 31).

Discussion

The reported generators for the visual EPN in the superior parietal lobule and inferior parietal lobule comprise brain areas (BA 5, BA 7, BA 40) that were shown to be activated during language processing, such as word comprehension, semantic categorization, speech perception and lexical decisions (Price et al., 1994; Seghier et al., 2004; Wilson et al., 2004; Bedny and Thompson-Schill, 2006; Chou et al., 2006).

Most studies that investigated the neural generators of the visual EPN, however, estimated the sources for the contrasts emotional minus neutral (as in our second approach) and estimated them in bilateral occipito-temporal and occipito-parietal cortical areas (Kissler et al., 2007; Schupp et al., 2007; Junghöfer et al., 2010; Schindler et al., 2015; Schindler and Kissler, 2016). Schacht and colleagues (2009b) followed a more descriptive approach, comparable to our first approach and showed a symmetric dipole pair in the fusiform gyrus for emotional words and exactly the same cortical brain area for neutral words. To our knowledge, only one study attempted to identify the neural generators for a possible auditory equivalent: In a typical EPN time frame (252–392 ms), Jaspers-Fayer and co-workers (2012) revealed activations in the superior parietal lobule (BA 7) for emotional words and syllables with emotional prosody. According to the authors, this area was shown to be a generator of the visual EPN and is associated with the control of selective attention to auditory stimuli. Therefore, the authors concluded that it might be a common neural generator of the EPN in both modalities and might once more suggest that selective attention is paid to emotionally salient stimuli. Interestingly, source localization for the auditory modality revealed comparable neural generators in the IPL and SPL, similar to the visual EPN in the present study, however with biggest activation cluster in the right hemisphere. This different lateralization across modalities is rather surprising, since language and speech understanding is commonly accepted to be a heavily left-lateralized function.

In general, the right hemisphere is thought to have a stronger involvement in imagination and intuitive, non-verbal behavior, whereas the left hemisphere is stronger involved in rational, verbal and analytical behavior. Emotional processing shows a comparable lateralization in terms of a differential involvement of the left and right amygdala (Markowitsch, 1998). The left amygdala is more strongly involved in emotion encoding with a higher affinity to language and detailed feature extraction, wehereas the right amygdala is more strongly involved in affective retrieval with a higher affinity to pictorial or image-related material. If extrapolate this findings for the amygdala to a general lateralization, the present results would allow the vague interpretation that spoken emotional words lead to a more pictorial imagination of the words as compared to the written words. According to the dual pathway model of auditory language comprehension syntactic and semantic information are primarily processed in a left-hemispheric temporo-frontal pathway, whereas intonational and prosodic information is processed in a right-hemispheric temporo-frontal pathway (Friederici and Alter, 2004). In the present study, spoken words were uttered with a neutral prosody to focus on the processing of emotional semantics only. However, it is conceivable that words with emotional content uttered with neutral prosody constitute a sort of incongruence between content and prosody. Therefore, there might be a stronger involvement of the right, prosody-processing hemisphere.

It is, furthermore, noteworthy that, the auditory activations were less strong and stable compared to the visual domain. The present results support the assumption that the superior parietal lobule is a common neural generator of attention effects to emotional stimuli being reflected in the EPN component across modalities.

Approach 2

Methods and Results

In a second step, we sought to compare sources across emotion categories, in order to make inferences on the brain networks involved in the emotion effects that were found between positive and neutral, or negative and neutral words in the two modalities. In order to do so, paired-samples T-Tests were computed in the source domain for the positive > neutral and negative > neutral contrasts (in the same frames as the separate tests: 234-350 ms for visual stimuli; 463-515 ms for auditory ones). Again, the plots only report voxels that were significantly more activated in response to the emotional words, as compared to the neutral ones (bullet-proof, corrected critical thresholds were calculated through 5000 randomizations in SnPM). The underlying neural sources corresponding to the time frames of the emotion effects were estimated with the statistical package implemented in sLORETA. . Results (summarized in Figure 3.6 and Table 3.3) showed that, for the visual domain, the main difference in activation between emotional and neutral words lay solely in the left occipito-temporal area, comprising areas stretching between the left cuneus and the left fusiform gyrus. Unfortunately, for the acoustic domain the analysis did not reveal any voxel in which the emotional and neutral words elicited a differential response, indirectly confirming that the signal in this condition might be less reliable or synchronized (across stimuli or participants).

Table 3.3

Source Localization with sLORETA of both experiments for emotion contrasts

	Visual Domain (234-350 ms)				Acoustic Domain (462-515 ms)			
	Critical T threshold	**Max T value**	**Lobe, Area**	**Cluster size (voxels)**	**Critical T threshold**	**Max T value**	**Lobe, Area**	**Cluster size (voxels)**
Positive > Neutral	3.38	4.48	Left MTG, BA 21	43	3.45	2.57 (n.s.)	Bilateral Frontopolar area, BA 8 and 9	/
		4.21	Left STG, Ba 22 and 41	31				
		4.17	Left ITG and Fusiform, BA 20	15				
Negative > Neutral	3.40	3.70	Left Occipital lobe, BA 18 and 19	11	3.42	1.68 (n.s.)	SFG, BA 10	/

Paired-samples T-Tests were computed in the source domain for the Positive > Neutral and Negative > Neutral contrasts. Reported are voxels that were significantly more activated in response to the emotional words, as compared to the neutral ones (bullet-proof, corrected critical thresholds were calculated through 5000 randomizations in SnPM). The underlying neural sources corresponding to the time windows of the emotion effects were estimated with the statistical package implemented in sLORETA.

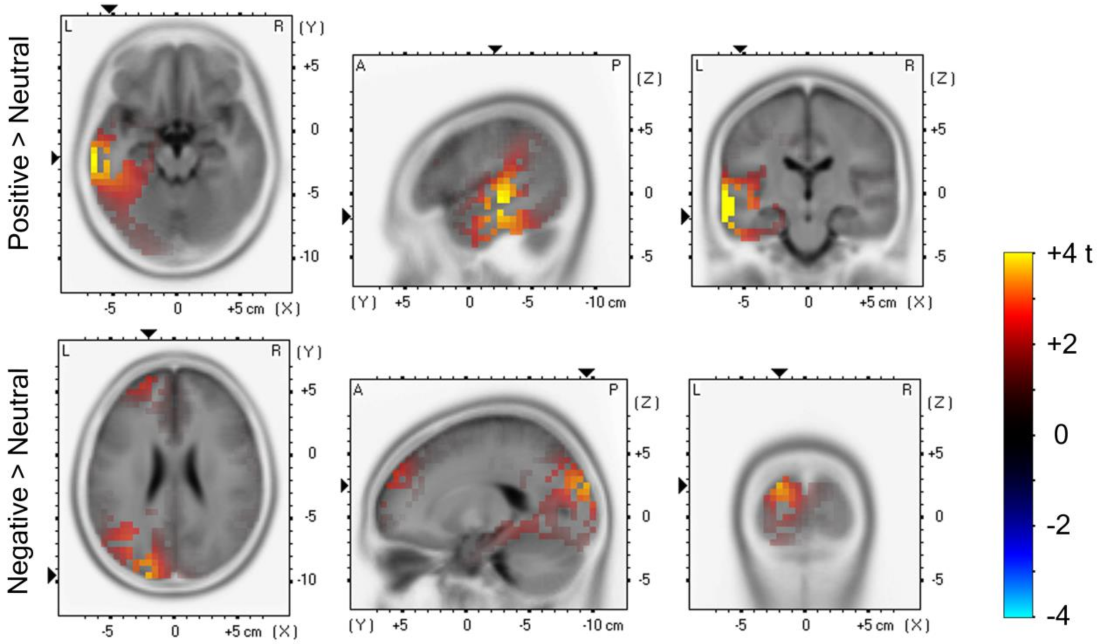

Figure 3.6 Source localization results for the emotional versus neutral contrasts. Depicted are voxels showing significantly higher activation in response to emotional as compared to neutral words. Sources are shown only for visual words. Source localization for auditory words did not reveal significant voxels. Main differences in activation between emotional and neutral words are shown in the left occipito-temporal area (left cuneus and the left fusiform gyrus).

Discussion

The present results for the contrasts between emotional and neutral stimuli revealed that the emotion effects on the scalp surface result from enhanced activation in left occipito-temporal areas to emotional compared to neutral stimuli, confirming the known neural generators for the visual EPN. Activation in BA 18 and 19 confirm an involvement of occipital, visual cortex areas, whereas the activations in the left fusiform area show an additional involvement of areas having language associated functions, such as selective processing of text and speech (Giraud, 2004; Vorobyev et al., 2004), word generation (Friedman et al., 1998) and semantic processing (McDermott et al., 2003; Chou et al., 2006). These results are in good accordance with studies investigating the sources of the visual EPN, estimating them in occipito-parietal and bilateral occipito-temporal cortical areas, as the fusiform gyrus (Kissler et al., 2007; Schupp et al., 2007; Junghöfer et al., 2010; Schindler et al., 2015; Schindler and Kissler, 2016).

For the auditory modality, in contrast, the differential response did not reveal significant results, pointing to a less synchronized and more widespread signal of the processing of spoken words with emotional content. Even though the difference map of positive compared to neutral words at the scalp level showed high similarity across both modalities, effect sizes for the scalp effects let assume a less reliable and synchronized effect in the auditory modality. Therefore, it seems conceivable that the second source analysis for the difference waves did not reveal a clear source for spoken words in the EPN time frame.

General Discussion

The present study aimed at a better understanding of emotion effects occurring on hearing spoken words with emotional meaning. In order to make a more profound statement to the question whether these emotion effects show similarity to the visual modality or even reflect equivalents to visual, emotion-related components, the study compared emotion effects occurring on hearing spoken words to those elicited by the same written words. To this end, two experiments were conducted, in which participants either read or listened to single nouns of positive, negative, or neutral content, while ERPs were recorded. The spoken words used in Experiment 1 were generated by a computerized female voice, while a human female voice was employed in Experiment 2. In both experiments, an occasional 1-back task was used in order to ensure that participants pay attention to the word stimuli.

In line with previous studies, written words of emotional content elicited an EPN consisting of an enhanced posterior negativity and an anterior positivity (Kissler et al., 2007; Schacht and Sommer, 2009b, 2009a), which were highly similar across the two experiments and showed the same latencies as well as a "positivity bias". Consistent with the literature, the findings indicate that words of emotional, and in particular of positive content, involuntarily capture the observer's attention as was shown for pictures (Junghöfer et al., 2001; Schupp et al., 2007) as well as words (Kissler et al., 2007; Kissler et al., 2009).

As the spoken modality of language should be of comparable, if not even of higher, relevance (see below) to a human subject, effects of emotional content were expected for spoken words as well. Indeed, the emotional content of spoken words similarly modulated ERPs in both experiments in the present study. Emotion-related ERP effects occurred at about 450 ms, that is about 200 ms later, compared to EPN effects in the visual modality. However, this delay is reasonable as full semantic information of spoken words develops over time as opposed to complete immediate presentation of visual information for written words (Bradley and Lang, 2000). Full stimulus information was available in average after 650 ms, clarifying the longer latencies for emotion effects in the auditory modality.

Beside this difference in latency, emotion effects in ERPs between the two modalities showed remarkable resemblances. The distribution of ERP differences between positive and neutral words that was observed here, deserves attention for several reasons. First of all, with its scalp distribution – posterior negativity and anterior positivity – it resembles the typical distribution of the visual EPN. This impression is verified by significant effects of emotion when applying exactly the same ROI to ERP analyses as in the visual modality.

With the aim of shedding light on the underlying neural mechanisms, we performed source localizations in order to estimate the neural generators of these activation patterns. Surprisingly, an at least partly overlapping pattern of neural generators was shown across modalities for all emotion categories lying within the superior and inferior parietal lobule. These activated areas were shown to be involved in language-related functions (Price et al., 1994; Seghier et al., 2004; Wilson et al., 2004; Bedny and Thompson-Schill, 2006; Chou et al., 2006). The SPL was, in fact, already suggested to be a common neural generator of the visual EPN and its equivalent in the auditory domain (Jaspers-Fayer et al., 2012), which is corroborated by the present results.

Interestingly, the typical activation in visual cortex areas was only revealed when analyzing the sources of the difference between emotional and neutral words for the visual domain. This would confirm the common assumptions concerning the functional significance of the visual EPN: In almost all studies reporting EPN effects to visually presented stimuli of emotional valence, including words (Kissler et al., 2009; Bayer and Schacht, 2014), emotional facial expressions (Schupp et al., 2004; Rellecke et al., 2011; Rellecke et al., 2012) and affective pictures (e.g., Schupp et al., 2004; Schupp et al., 2012), EPN findings have been suggested to reflect enhanced allocation of sensory resources in the visual cortex, resulting from re-entrant activation of particularly extra-striate areas by cortico-amygdaloid structures (Vuilleumier et al., 2004; Kissler et al., 2007). The present study revealed such a boost in the visual cortex for emotional in contrast to neutral stimuli on investigating the sources of the difference waves (emotional minus neutral). For the auditory domain, this analysis did not reveal significant results, probably due to less synchronization of activation across participants. Another characteristic that both components have clearly in common is the positivity bias: ERP effects of emotional content consisted of enhanced posterior negativities and fronto-central positivities that were mostly pronounced for positive words in both experiments. The present findings, therefore, provide evidence that the bias for positive emotional valence in word processing is not restricted to or specific for reading. This is in line with Rohr and Abdel Rahman (2015), who showed a positivity bias for spoken words with more pronounced emotions effects for positive stimuli that were, however, restricted to the communicative situation.

All in all, it was possible to demonstrate that, in addition to established effects of emotional content in the visual domain, effects also exist for spoken emotional words. These effects were revealed at the same latencies - invariantly of the speaker's voice being human or artificial. Overall, emotion effects for spoken word processing had longer latencies and were weaker compared to the visual emotion effects. These findings might be explained by a stronger jitter in the spoken stimulus material caused by the incremental nature of this type of stimuli. The effects showed clear resemblances to the visual EPN as well as shared sources with their visual counterparts.

4 Comparison of Emotion Effects for Spoken Words between Study 1 and Study 2

The two studies of my dissertation were aiming at investigating effects of emotional content on event-related potentials occurring to spoken words and the comparison with visual emotion effects. In both studies, the same words spoken by a female human voice and in Experiment 2 of the second study additionally by a computer-generated voice were presented to the participants. Surprisingly, as can be seen in Figure 4.1 (panel A), emotion effects across both studies differed nevertheless in several aspects: First, in Study 1, effects of emotional content of spoken words started roughly 100 ms earlier than in the second study. Second, the topography of the effects in Study 1 showed less similarity to the visual EPN as in Study 2 and is rather comparable to a N400 component. Last, emotion effects in Study 1 were solely driven by negative emotional content; in Study 2, effects were mainly driven by positive content. These differences were surprising as the high concordance of both studies in terms of stimuli and study design is apparent. The question arose whether the differences between both studies, more precisely the volume level manipulation or the additional visual presentation constitute the reason for the differences in emotion effects. If an additional visual presentation caused the differences in the emotion effects this would imply that a priorly formed visual percept of the words interacts with the auditory processing of emotional content. Therefore, this would speak for a close connection of language processing in both modalities at level of emotional content. This chapter seeks to compare both studies and the corresponding emotion effects in more detail and describes subsequent investigations made to elucidate the origin of the revealed differences.

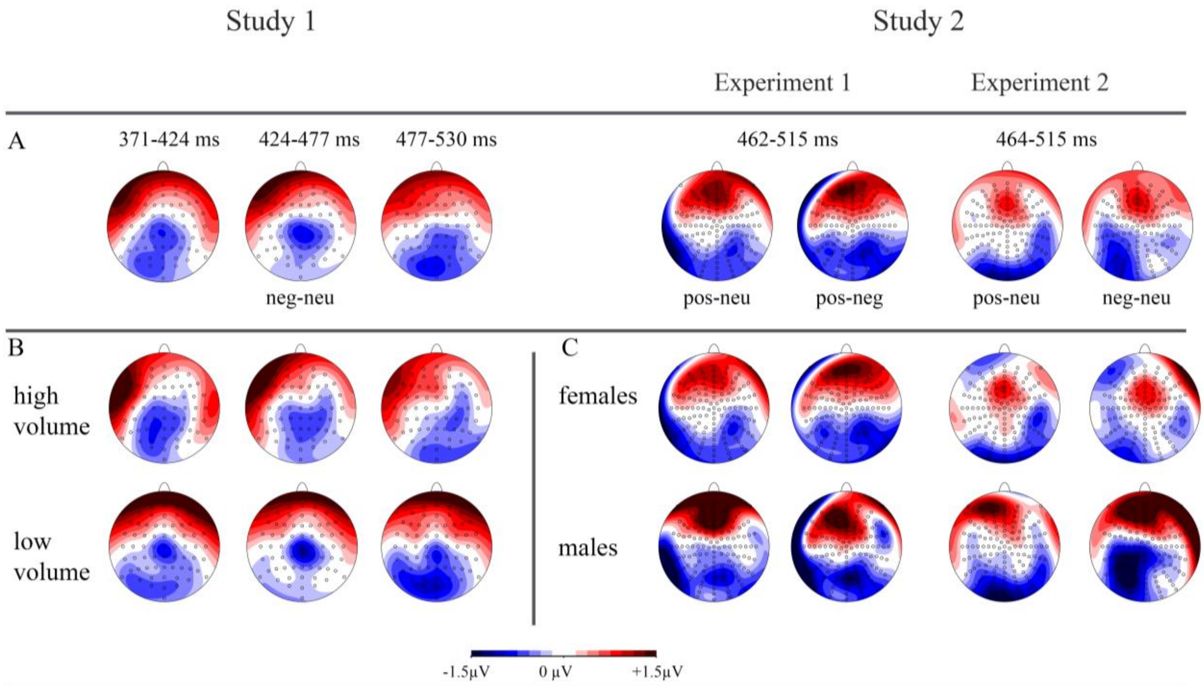

Figure 4.1 Comparison of emotion effects in both studies.
(A) Comparison of revealed emotion effects in Study 1 and Study 2. The left panel shows significant emotion effects of Study 1 as scalp distribution of the differences between negative and neutral words in the depicted time frames. The right panel shows significant emotion effects of Study 2 as differences between indicated emotion categories in depicted time frames. Panel **(B)** depicts emotion effects of Study 1 in the same time frames and for the same contrast (negative-neutral) as in Panel **(A)** separated for high volume and low volume words.
Panel **(C)** depicts emotion effects of Study 2 in the same time frames and for the same contrasts as in the Panel **(A)**, separated for female and male participants.

General similarities and differences across both studies

Stimuli

Both studies were conducted with spoken word material that was initially selected in the visual modality with several important linguistic parameters carefully controlled across the positive, neutral and negative category (Bayer et al., 2012a). In the visual modality, these stimulus words led to effects on the emotion-related EPN and LPC component (Bayer et al., 2012a). Only two nouns of the original stimulus material had to be replaced because of their ambiguous phonology. In Study 1 and Experiment 1 of the second study, exactly the same stimulus material was used, whereas in Experiment 2 the same words were produced by a computer voice-reader software.

Task

Next to the same stimulus material, also the task was the same in both studies. In the one-back task, participants were asked whether the presented word was the same as the immediately preceding word or a different one. This query was interspersed at the same frequency in both studies. The test word presented in the one-back task trials was visual in study 1 and Experiment 2 of study 2, since only the seventy-two stimulus words were recorded by the female speaker. InExperiment 1 of Study 2 1-back items were matched in modality to the one of the given block (auditory or visual). In Study 2, participants received a feedback about the correctness of their answer whereas in Study 1 no feedback was given.

Participants

In Study 2, seven male and eighteen female participants took part in Experiment 1 with a mean age of 22.4 years (SD = 3.4) and eight male and fourteen female participants with a mean age of 21.5 years (SD = 2.5) in Experiment 2. In Study 1, the sample was slightly bigger including twenty-nine participants with a mean age of 23.7 years (SD = 2.8). In contrast to Study 2, only female participants took part in the first study.

Design

The most obvious differences between both studies became visible in the study design. In Study 1, stimulus words were presented at two different volume levels within one block in randomized order whereas words were presented with a stable volume level in Study 2. Furthermore, the second study included an additional visual presentation block. As block order was counterbalanced across participants, half of them have read the words before hearing them.

Another difference between both studies was the number of electrodes. In Study 1, ERPs were recorded from 64 electrodes, whereas in Study 2 they were recorded from 128 electrodes.

Further analysis for elucidating differences in emotion effects across both studies

Even though at first appearance both studies show high concordance, on closer inspection some structural differences were listed above. These differences were followed up where possible in further analyses to exclude a potential influence of these factors on the emotion effects and their differences.

Topography differences

Volume level manipulation

As discussed before, the volume level manipulation was one of the most obvious differences between the two studies. In the first study, the stimuli were played to the participants at two different volume levels where the words were randomly assigned to one of the two volume levels. One could assume that the random presentation at two different volume levels led to a stronger focus on the volume level and thereby probably withdrew attention from the emotional content of the word. This might have led to an altered processing of the stimuli and thus to differences in the ERP results.

However, as was revealed in Study 1, volume level was not interacting with the factor emotion. This would argue against volume level manipulation being the cause for the differences in emotion effects. Nevertheless, since it was one of the most obvious differences between the two studies, visual inspection and topographic comparisons were calculated comparing high volume and low volume words to elucidate whether differences are shown between the two categories. Mean amplitudes in the relevant time windows were normalized within each condition (high and low volume) and participant by dividing them by the GFP. Normalized amplitudes were investigated with repeated-measures ANOVA with the factors electrode x volume level. Furthermore, it was of interest if one category shows higher similarity to the emotion effects in Study 2.

Figure 4.1 (panel B) shows the difference maps for negative compared to neutral words separately for the high and low volume condition. Overall, difference maps for both volume levels show a comparable distribution of a frontal positivity and a central to posterior negativity. This observation was implicitly approved by the topographical comparisons, revealing no significant differences between the maps for high and low volume words and none of the investigated time windows, all $Fs(63, 1764) < 1.7$. It is visible especially in the first two time frames that there is a comparatively isolated negativity around central electrodes for low volume words. The negativity for high volume words does spread from posterior to central electrodes as well, however being less spotty. Nevertheless, none of the volume level conditions shows clear similarities with the EPN-like distribution of Study 2.

Additional visual presentation in Study 2

Apart from the missing volume level manipulation in Study 2, words were additionally presented in their written modality. The subsequent visual and auditory presentation was realized in a block design and the order of blocks was counterbalanced across participants. Consequently, for half of the participants the words were unknown when hearing them in the auditory block, whereas the other half had already read them before. Possibly, the visual presentation of the words before the auditory presentation simply have led to some kind of repetition effect and thereby to a differential processing of the auditory stimuli.Therefore, it was investigated whether the block order (first visual then auditory block or vice versa) influenced the event-related potentials occurring to the spoken words. If there were an influence of block order on the auditory-evoked components, it would be conceivable that the missing visual presentation in Study 1 would have caused the differences in emotion effects across the studies. Due to a lack of power when splitting the participants into two groups, data of Experiment 1 and 2 were collapsed for the analysis in the time frame 464-515 ms. Again, repeated-measures ANOVA were calculated in the respective time frame of the auditory emotion effect, including the within subject factors emotion (3- positive, neutral, negative), cluster (16 á 7 electrodes) and block order as between subject factor. This analysis revealed no effect of block order on the emotion effect in the auditory modality, $F(1,45) < 1$. This indicates that it did not make any difference whether participants had read the words before hearing them. Therefore, the missing visual presentation of the words did not seem to lead to differences in emotion effects across both studies.

Gender effects

In Study 1, data was collected solely from female participants, whereas in Study 2, both male and female participants took part. Even though, there was no specific reason for that, it would still be possible that men and women process the words differently. Therefore, it was investigated whether in Study 2, ERPs to spoken words show differences for male and female participants. Again, data of both experiments were collapsed for better statistical power. Gender was included as between-subject factor in the overall ANOVA.

The analysis revealed a significant effect of gender on the overall activity in the time frame 464-515ms, $F(1,45) = 5.228$, $p < .05$, $\eta_p^2 = .104$. However, no interaction of the factor gender with any other factor (cluster, electrode or emotion) was revealed. Only a trend for significance was discernible for the interaction gender x cluster, $F(15,675) = 2.155$, $p = .096$, $\eta_p^2 = .046$. Post-hoc analysis in the anterior and posterior ROI defined for the emotion effect only revealed trends for an influence of gender, $F(1,45) = 3.137$, $p = .083$, $\eta_p^2 = .065$ and $F(1,45) = 2.585$, $p = .115$, $\eta_p^2 = .054$, respectively. As becomes visible also from Figure 4.1 (panel C), the effect of the participants' gender was mainly driven by Experiment 2 (human voice), showing higher discrepancies between female and male difference maps. For female participants, difference maps show an isolated positivity at fronto-central electrodes, whereas the positivity for male participants is broader and spans further towards frontal electrodes. The complementing negativity for male participants is shown highest at posterior electrodes for positive words and spans further towards parietal electrodes, whereas for female participants the negativity is less distinct and even spreads towards frontal electrode sites. Overall, for male participants higher amplitudes are shown as compared to female participants. In addition, it was confirmed that Experiment 2 mainly drove the gender effect by overall ANOVAs calculated separately for both experiments. For Experiment 2, this ANOVA showed a trend towards significance of gender as between-subject factor for Experiment 2, $F(1,20) = 3.364$, $p = .082$, $\eta_p{}^2 = .144$, but a non-significant result for Experiment 1, $F(1,23) = 1.285$, $p = .269$, $\eta_p{}^2 = .053$.

Latency differences

In Study 1, emotion effects were revealed starting 370 ms after word onset and lasting up to 530 ms. In Study 2, emotion effects showed longer latencies and started only 460 ms after word onset. The last time frame in Study 1 (477-530 ms) shows the highest resemblance with the EPN-like emotion effect of Study 2. This was confirmed by analyzing amplitude differences between 477-530 ms in a comparable anterior and posterior ROI as in Study 2. For both ROIs emotion effects were revealed in anterior and posterior ROI, $F(2,56) = 3.589$, $p < .05$, $\eta_p{}^2 = .114$. and $F(2,56) = 4.344$, $p < .05$, $\eta_p{}^2 = .134$, respectively. The earlier onset of the emotion effect together with the different distribution at scalp topography level might lead one to assume that in Study 1 an earlier effect (e.g. N400-like) is interfering with or preceding the EPN-like effect. Figure 4.2 shows again the scalp distribution of the difference between negative and neutral words in the three consecutive time frames between 371 and 530 ms of Study 1. In comparison, the figure shows the same contrast of negative minus neutral words for both experiments of Study 2 in the same time frames. For Experiment 1, the distribution of difference waves at scalp level shows no clearly delimited distribution of positivities and negativities for this contrast. The scalp topography for Experiment 2, however, shows a frontal positivity and a negativity at parieto-occipital electrodes. This distribution shows high similarity to that of Study 1, being slightly more lateralized to the left. To verify whether this time frame of Study 2 shows a resembling emotion effect, a repeated-measures ANOVA was calculated in a parieto-occipital ROI comparable to that of Study 1. No significant difference between emotion categories in any of the investigated time frames of Experiment 2 was revealed, all $F(2,42) < 1.9$.

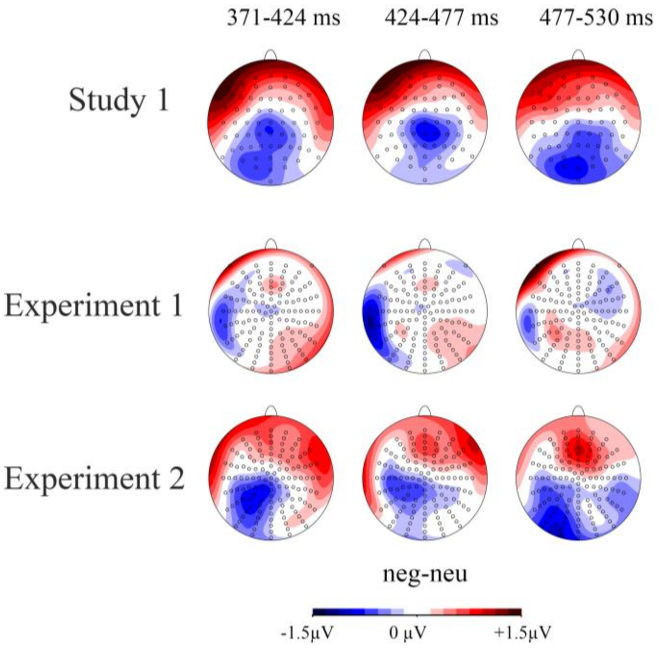

Figure 4.2 Comparison of emotion effects of Study 1 with the same contrasts at the same latencies of Study 2. The upper panel depicts significant emotion effects as difference maps between negative and neutral stimuli revealed In Study 1. The lower two panels depict the same contrasts for Experiment 1 and 2 of Study 2 at the same latencies as emotion effects of Study 1.

Discussion

This chapter of my dissertation was aiming at getting a better understanding of the existing differences in emotion effects across the two studies. Additional visual inspections and statistical analysis of the data were done to find the underlying cause of the differences in terms of latencies and scalp distributions of the corresponding difference waves. First, I focused on the differences in scalp distributions of the emotion effects. The participants' gender was investigated as a possible influence in order to clarify whether the restriction to female participants in Study 1 could have caused differences in the neural processing of emotional words. An influence of gender in the time frame of the emotion effect in Study 2 was actually detected, especially driven by the second experiment. However, this effect was neither interacting with the factor emotion nor detected in the ROIs of the emotion effects. Sex differences on event-related potentials are evident in the literature in several studies:

Electrophysiological correlates of initial lexical-semantic access were shown to be similar in men and women, whereas higher order semantic processing differed between the sexes, being evident in earlier and longer lasting context effects on the N400 component for women (Wirth et al., 2007). Shorter latencies and higher amplitudes of N400 effects in women were shown in auditory semantic processing as well, whereas men showed aufmented LPC effects (Daltrozzo et al., 2007). These results suggest a dissociation between men and women during semantic (auditory) processing; therefore, it is possible to find a dissociation in processing of emotional semantics as well. However, the present results show stronger effects for men compared to women, which is in contrast to the results of the above reported studies. Importantly, a further look at the scalp topographies of emotion effects for female participants revealed no higher similarity to emotion effects in Study 1. Therefore, it seems improbable that the additional male participants in Study 2 led to the differences in scalp topographies of the emotion effects across studies. Nevertheless, it may be noted that the processing of spoken words (at least in this specific time frames) seem to differ between female and male participants.

In addition, it was investigated whether the additional visual presentation of the stimulus words in Study 2 influenced topographies of the emotion effects. It would be conceivable that a previously built visual percept or the higher number of repetitions of each word led to a different processing of the auditory words. However, the results revealed no influence of block order, which means by implication it made no difference whether participants had read the words before the auditory presentation or not. Consequently, in my opinion, it is unlikely that the differences in emotion effects across both studies were caused by the additional presentation of the same words in the visual domain.

Even though volume level was not interacting with the factor emotion in Study 1, a depiction of the difference maps of negative and neutral words separated by volume level condition showed slight differences between high and low volume condition. This might indicate that, even though not found in the ANOVA across all electrodes, volume level might to some extent influence emotional processing, leading to slight differences, specifically in a central negativity.

In addition to the differences at scalp topography level, emotion effects differed in their latencies across both studies. Emotion effects in Study 1 were revealed in three consecutive time frames between 370 and 530 ms, whereas in Study 2 emotion effects only occurred roughly between 460 and 520 ms. A visual inspection of the data of Study 2 in the same (earlier) time frames as in Study 1 (with the same contrast - negative minus neutral) revealed topographies with resemblances to Study 1 for experiment 2 (Figure 4.2). This resemblance appears to be much bigger than the initial comparison of the emotion effect across both studies in the different time frames (Figure 4.1, Panel A). This observation allows the speculation that the effects not simply differ in terms of latencies and topographies, but that Study 1 shows an additional emotion effect, with a more N400-like topography as discussed before in chapter 2. This in turn would explain the latency- as well as the scalp topography differences. In the latest time frame of Study 1 (477-530 ms), this N400-like effect seems to merge into an EPN-like topography. When only comparing this last time frame with the emotion effects of Study 2, it becomes apparent that these overlapping time frames in both studies show higher concordance in terms of topographies than the two earlier time frames.

When assuming that the topography and latency differences result from an additional preceding emotion effect in Study 1, the question remains open what actually caused such a supplemental effect. Furthermore, the reason for the valence differences of the emotion effects were not answered by now, as in Study 1, emotion effects were driven by negative content only, whereas in Study 2, effects were mainly resulting from positive valence. In my opinion, with the present data it was not possible to elucidate where this difference was resulting from and future research is needed to answer the open questions.

All in all, the findings concerning the cause for the differences across both studies appear rather heterogeneous and do not point to a clear result. When taking into account all analysis described above, I would argue that in Study 1 the emotion effect of Study 2 is preceeded by an additional emotion effect. Unfortunately, with the additional analysis in this chapter it was not possible to reveal the origin of this additional emotion effect.

When considering the earlier effects in Study 1 as being N400-like effects, it is conceivable that for negative words the presentation at different volume levels might have matched the expectation of the participants less than for neutral words, as the N400 is often occurring to expectancy violations. Negative stimuli comprised words related to both sadness and anger. Sadness is an emotion that is most commonly portrayed less intense and in terms of vocal utterances with emotional intonation linked to smaller amplitudes (Scherer, 2003). Contrary to this, anger is an intense emotion and connected to higher amplitudes. This close association of different emotions with different amplitude levels might generalize to spoken utterances with emotional content as well. Therefore, a presentation of words of sad content with a high volume might have led to violations of expectancies and therefore to N400 effects. The same might be true for words of anger-related content presented at low volume level. However, a closer look on the negative stimulus words revealed that only fourteen of the twenty-four stimulus words were clearly assignable to one of the two categories, resulting in a lack of power. The calculated difference maps of low volume angry words minus high volume angry words and high volume sad words minus low volume sad words did not show a N400-like topography. Nevertheless it would still be possible that the volume level manipulation was the causal agent for the differences. Not knowing whether to prepare for a high or low volume level might have changed the whole context of the experiment and in turn led to different emotion effects as well. However, future research is needed to finally calrify the reason for the different emotion effects to spoken words.

5 General Discussion

The main aim of the present dissertation was to investigate effects of emotional content on auditory event-related potentials evoked by single spoken words. For written word processing, several studies had investigated the boundary conditions and functional locus of emotion effects on event-related potentials. In contrast to this, for the auditory modality almost no study had investigated the effect of emotional content on spoken word processing. Thus, this work was designed to investigate whether emotional content does modulate auditory-evoked potentials reliably and how attention driven by physical stimulus properties and social significance of the speaker might interact with these effects. Within this it was of particular interest, whether the effects in the auditory modality show similarities to emotion effects to words in the written modality.

For this purpose, in two studies the processing of single spoken nouns uttered in a neutral prosody and their possible modulation by emotional content was investigated to their full extent. First of all, effects of emotional meaning on event-related potentials were evident in both studies. Accordingly, it can be concluded that the processing of single spoken words is modulated by emotional content of a word, as was shown for visual word processing in numerous studies. The question remains whether these emotion effects show similarities to the visual counterparts or whether they occur in a different way.

In Study 2, emotional meaning modulated ERP amplitudes roughly between 460 and 520 ms. This effect was stable and occurred in both experiments unaffected by the social significance of the speaker's voice. Even though segmentation borders were revealed independently of each other, effects occurred at almost exactly the same latencies in both experiments, thus confirming the reliability of the effect. It should be noted that, overall, effects of emotional content on auditory-evoked potentials are weaker than in the visual modality, which is reflected both in effects at scalp topography level and in analyzed sources. The effects were mainly driven by positive valence compared to neutral content; however, negative content elicited a similar effect (Experiment 2). From a merely observational point of view, this component shows quite a high resemblance with the EPN known for visual emotional stimuli, more precisely a posterior negativity with its counterpart, an anterior positivity. However, it has to be emphasized that a resemblance is only revealed for the difference between emotional and neutral stimuli. The underlying topographies for the single emotion categories look highly different across both modalities. Therefore, the question has to be posed whether the purely observational resemblance is sufficient to determine this component as being the counterpart to the EPN for written words.

To my knowledge, the only evidence for the existence of an “auditory EPN” originates from two studies carried out by the same group (Mittermeier et al., 2011; Jaspers-Fayer et al., 2012). In my opinion, there is some vagueness concerning their approach and their motives for defining their effects as “auditory EPN”. First, they do not compare the processing of emotional content to processing of neutral content but the processing of emotional syllables (uttered with emotional prosody) or words of emotional content to the processing of neutral tones. Therefore, these effects might reflect more general differences in stimulus processing than solely differences in emotional meaning. Furthermore, they do not explain their motives for defining this component as an equivalent to the visual EPN. Mittermeier and colleagues define the highest negative value at three parietal electrodes (P3, Pz and P4) as EPN “peaks” and Jaspers-Fayer and colleagues investigate amplitudes at one single parietal electrode (Pz). Finally, the experimental set up used by both groups is identical, but the allegedly identical “auditory EPN” occurs in rather different time frames. Jaspers-Fayer and colleagues found effects between 132 to 156 ms and from 252 to 392 ms after word onset, whereas Mittermeier and colleagues revealed effects 150 and 190 ms post-stimulus (stimulus durations: tones and syllables: 250 ms; words: 500 ms).

Certainly, there are several criteria for defining a component as an equivalent to a known ERP as for example comparable latencies, comparable scalp topographies, having the same functional locus or the same neural generators. In terms of latencies, emotion effects found in Study 2 differed by about 200 ms from the onset of the visual EPN. However, the same latency for an auditory equivalent would have been rather surprising due to the incremental nature of auditory stimuli. On average, the presented words have a duration of 650 ms, hence it takes some time until the content of the words builds up, whereas in the visual modality complete information is available at onset of stimulus presentation. Nevertheless, Marinkovic and colleagues (2003) showed a converging of activation in supramodal networks for semantic access for visual and auditory word processing around the same time point (~ 230 ms after word onset). However, I want to bringt into focus that auditory words in their study were shorter than the written words (on average 0.7 syllables shorter).

Some aspects indicate a comparable functional locus of the reported effects to the visual EPN as well: First, both are evoked by the same words within the same group of participants. Additionally, both effects show a stronger impact of positive valence than of negative valence. Furthermore, like the visual EPN, also the auditory emotion effect presumably occurs at a post-lexical processing stage. However, for the auditory modality, this is only an assumption, since lexical access is assumed to have taken place in order for emotion effects to occur. Finally, further investigations would be needed to determine whether both components have a comparable functional locus, for example to answer the question if the auditory effect is shown for different stimulus domains as it was shown for the visual EPN.

In addition to a comparable functional locus, a counterpart in a different modality would furthermore be reflected by comparable neural generators. Since the visual EPN is known to reflect an enhanced allocation of sensory resources in the visual cortex (e.g. Schupp et al., 2004; Kissler et al., 2009; Bayer and Schacht, 2014), for an auditory equivalent it would be expected to find an enhanced activation of auditory areas, respectively. Investigating the underlying neural sources revealed activity in at least partly overlapping areas in the inferior and superior parietal lobule, which are known to have language-related functions. The SPL has in fact already been suggested to be a common neural generator of the visual EPN and its equivalent in the auditory domain (Jaspers-Fayer et al. 2012), which would be corroborated by the present results. However, it is important to note that these results only point towards an underlying basal activity in comparable brain areas, however, not to a comparable boost of brain activity driven by emotional content. Since the resemblance of emotion effects in both modalities is found in the difference waves between emotional and neutral stimuli, the analysis of the sources of the difference between emotional and neutral words would correspondingly be an important comparison. This analysis revealed stronger activation in the visual cortex for the emotional written words compared to neutral words, confirming the classical assumption of the visual EPN reflecting enhanced allocation of sensory resources in the visual cortex across different visual stimulus domains (e.g. Schupp et al., 2004; Kissler et al., 2009; Bayer and Schacht, 2014). Unfortunately, the same investigation of the sources of the contrasts revealed no significant results for auditory stimuli and the non-significant results rather point towards prefrontal areas. Overall, source localization results do not support the hypothesis of the effects reflecting an “auditory EPN” any further.

The reason that primarily led to this classification as “auditory EPN” is the topography of the difference waves showing a classical EPN-distribution with a posterior negativity and its counterpart, an anterior positivity. Even though showing close resemblances across the auditory and visual modality, there are two important points that should be noted: First, for most time frames, effects of emotional compared to neutral content were revealed in both, posterior and anterior regions. For the visual modality, effects were strongly driven by the posterior negativity whereas auditory emotion effects relied stronger on the anterior positivity. Second, the resemblance was only revealed when looking at difference waves for emotional minus neutral stimuli. EPN effects are usually only depicted as difference maps in the literature. However, when considering the underlying topographies for the single emotion categories, effects across modalities do not resemble each other at all. For visual stimuli, the underlying topographies show frontal negativities and posterior positivities, but centro-parietal negativities and frontal positivities for auditory stimuli. Accordingly, the posterior negativity in the visual modality results from higher posterior positivity by neutral words, whereas in the auditory modality, there seems to be a stronger anterior positivity and posterior negativity to emotional as compared to neutral words. This demonstrates that different topographies can lead to comparable difference maps, which not necessarily implies they reflect the same processes. The fact that the EPN is visible as a difference component therefore causes additional difficulties in investigating whether a specific component has an equivalent in a different domain.

Summing up the findings discussed above, it can be concluded that the EPN occurring to emotional visual stimuli in Study 2 reflects enhanced activity in the visual cortex. Even though basic activity during the time frame of the visual EPN and the auditory emotion effect is shown in at least partly overlapping brain areas, the activity at scalp topography level shows a highly different pattern. In the difference waves for emotional minus neutral stimuli, these different scalp topographies lead to a similarly distributed emotion effect, however, in my opinion, this is not reflecting the occurrence of the same component for both modalities. In the following, I will call the reported emotion effect evoked by spoken words with the frontal positivity and posterior negativity **auditory emotion-related component (AEC)**. The AEC occurred in both experiments Study 2, even though acoustic stimuli differed from each other across the two experiments. Overall, there were no differences in acoustic parameters between the groups of both experiments, which, nevertheless, does not exclude different acoustic profiles at single word level, e.g. different durations or fundamental frequencies. Furthermore, due to the different speakers' voice, the auditory stimuli differed in terms of their social significance to human listeners. Nevertheless, the auditory emotion-related component was revealed in both experiments at the same latencies reinforcing a stability and reliability of this effect.

An attempt was made in Study 2 to elucidate the sources of the AEC, in order to make a statement about the neural pathway of emotional auditory word processing and compare it to the written modality. According to Hickok and Poeppel (2007), the cortical organization of speech perception is based on a dual-stream model, composed of the ventral and the dorsal stream. The ventral stream is built by superior and middle portions of the temporal lobe and involved in the processing of speech signals for comprehension (speech recognition). The dorsal stream is built by the posterior frontal lobe and the posterior dorsal-most part of the temporal lobe and the parietal operculum. It is involved in translating acoustic speech signals into articulatory representations and builds an interface with the motor system (speech production). In accordance with this, neural generators of the AEC would be expected to be found in temporal areas along the ventral stream (rather than the dorsal stream) since the task did not require speech production but only speech recognition. Highest activations during the time window of the AEC, however, were shown in parietal areas as the inferior parietal lobule and the superior parietal lobule with a dominance of the right hemisphere. However, when taking into account as well the timing of speech processing, it seems likely that the AEC is occurring only later in time. Marinkovic and colleagues (2003) corroborated the dual stream model by Hickok and Poeppel showing auditory word processing to start in superior temporal areas (~55 ms), progressing to the preisylvian cortex (~100 ms) and proceeding along the ventral stream into anterior and lateral areas of the superior temporal gyrus (~250-300ms). After 300 ms activity w shown in a supramodal semantic network as in temporal and inferior prefrontal regions. Even though, activity in parietal regions was not shown by Marinkovic and colleagues, the IPL and SPL were shown before to have language processing-related functions and might be part of the semantic network.

In Study 1, effects of emotional compared to neutral content were revealed as well, mainly driven by negative content. These effects differed in terms of topography as well as latency from the auditory emotion-related component of Study 2. As discussed in chapter 4, these differences are most likely explainable by a preceding emotion effect prior to the auditory emotion-related component. Presumably, the AEC in Study 1 occurred between ~470 and 530 ms, thus at similar latencies as in Study 2. Approximately, between 370 and 470 ms, the AEC was preceded by an emotion effect with a N400-like topography. In comparable paradigms, such as visual congruency or syntactic and semantic violation paradigms, the N400 was shown to exist similarly for the auditory modality (e.g. Hahne and Friederici, 2002; Schirmer and Kotz, 2003; Wicha et al., 2003; Balconi and Pozzoli, 2005; Lück et al., 2006; Erlbeck et al., 2014). Furthermore, Marinkovic and colleagues (2003) have demonstrated a transition from modality specific-streams to supramodal networks for semantic access being particularly evident for the N400 component. However, evidence for a close relation of the N400 with emotional processing is scarce and was shown only in a few studies for visual stimuli (Kanske and Kotz, 2007; Herbert et al., 2008; Holt et al., 2009). Additionally, it has to be mentioned that N400 effects recorded with comparable setups (mainly average reverence) usually occur with a more central maximum. At the end, the question remains what caused the additional N400 effect and why it was absent in Study 2. This question needs to be adressed in future studies.

Conclusion

From the outcome of the present investigations, it is possible to conclude that the processing of spoken words is modulated by emotional content. This modulation is reflected in an anterior positivity and posterior negativity around 460 to 520 ms after word onset, referred to as an auditory emotion-related component. The AEC occurred in response to both the human voice and the artificial computer voice, and was not affected by the social significance of the speaker's voice. Furthermore, it was present even though the context of experiment design was presumably changed by a presentation at different volume levels. At the difference level of emotional and neutral content, the AEC bears some resemblance to the distribution of the visual EPN; however, mainly due to the different underlying topographies, this component is not considered to be an equivalent to the visual EPN. Furthermore, for the visual modality, the EPN is assumed to reflect enhanced allocation of visual sensory resources. For this reason, a comparable topography for an "auditory EPN" is unlikely to be found. Since effects were revealed at both frontal and posterior electrode sites, the component is not classified as positivity or negativity but more general as emotion-related component. If anything, the effect could be better classified as frontal positivity than posterior negativity, since effects were stronger in frontal regions at scalp topography level and also source localizations hint towards an involvement of prefrontal areas, however not reaching significance. Concerning a comparison of the underlying neural network during the processing of content in spoken and written words, it can be concluded that the processing activates brain areas which are at least partly shared in the SPL and IPL. However, when it comes to the processing of emotional content compared to neutral content, the analysis of neural sources for visual words show an enhancement of activity in the visual cortex whereas for the auditory words this analysis did not reveal a satisfying outcome. Nevertheless, also based on the different scalp topographies for single emotion categories in both modalities, it can be hypothesized that emotional content in both domains interacts differently with neural semantic processing.

Limitations and Future Directions

Even though the experimental results point towards a strong reliability of the auditory emotion-related component, nevertheless some limitations were identified over time. In the following, problematic aspects and directions for future research are discussed.

A first limitation of the present study is the stimulus material, which was taken from a former study (Bayer et al., 2012a). For the visual modality, these words were shown to elicit typical emotion-related ERP components. Furthermore, stimulus categories were controlled for several important linguistic parameters, such as word frequency, number of letters and syllables and imageability ratings. On average, stimulus words were 2.03 syllables long, however the shortest word was one syllable and the longest four syllables. For a visual presentation of words, such variability in length is not problematic due to the instantaneous presentation of stimulus information. However, for auditory stimuli, a stronger limitation of word length would be necessary, since here stimulus information is presented serially. Consequently, for auditory words, the point in time where enough information is available for uniquely identifying the right word (uniqueness point) differs. Presumably, this variability in the length of stimulus words led to a poor synchronicity of the processing of content and thus to a reduced synchronicity of emotion effects. It is conceivable that this low synchronicity led to the weaker effects at scalp topography level as well as in the investigated sources. Possibly, the comparison of neural activity for the contrasts emotional minus neutral would have led to significant results if the stimulus material would have allowed for a better synchronicity of the neural response. Therefore, I want to emphasize at this point the importance of strictly controlled stimulus material for the auditory modality of word processing.

Another approach to achieve a better synchronization of auditory ERPs would be a segmentation of data according to the uniqueness points or recognition points of the words instead of word onset. Computed by reference to a phonetic dictionary, the uniqueness point is the earliest moment at which a word can be uniquely distinguished from all others in terms of a sequential phoneme-by-phoneme comparison. There is empirical evidence that the uniqueness point bears close relationship to the recognition point, that moment in time at which a word can be recognized from the acoustic input (cf. Marcus and Frauenfelder, 2007). Since the uniqueness points of each word differed, consequently their recognition process differed as well and might have taken longer for some words than for others, which in turn led to a poor synchronization of the electrophysiological signal. Unfortunately, the Berlin Affective Wordlist (Võ et al., 2009) from which the words were selected did not contain information about the uniqueness points. Attempts to find the missing information in other databases was not successful either. Furthermore, a segmentation according to uniqueness point would presumably not be sufficient enough, since it is not clear whether it can be equated with the recognition point of the words or whether the latter is affected by factors as word frequencies or orthographical neighbors. In order to reduce the temporal jitter in the data to a minimum, a segmentation according to the recognition point would be the most suitable approach. However, the determination of the recognition points of the present spoken word material would require a complex experimental investigation, which was not realized during this dissertation. In addition to that, the initial idea was to directly compare the effects of emotional content in the spoken modality to those for written words. For "normal" visual word presentation (as opposed to rapid serial visual presentation RSVP for example), no recognition point exists since whole information is available upon word onset. In order to determine when, in comparison to visual word processing, emotion effects appear in the auditory domain, it was therefore necessary to investigate them based on word onset. Continued research on emotion effects in spoken word processing should consider to select stimulus material from databases, which contain the uniqueness points of the words as well or determine the recognition points in a separate experiment with a successive presentation of word fragments.

The present dissertation was specifically aiming on investigating effects of emotional content on spoken word processing isolated from (emotional) prosody. This was done in order to be able to directly compare the effects occurring in the spoken modality to the known effects for written words, which do not contain this second information channel. However, it has to be mentioned that these two channels (semantic information and prosody) are intertwined and show complex interactions (e.g. Wambacq and Jerger, 2004; Schwartz and Pell, 2012). It has to be taking into account that words with emotional connotation but spoken with a neutral prosody might represent a sort of discrepancy or incongruence, possibly eliciting a conflicting processing. Therefore, future studies should include different emotion categories (positive, neutral, negative) at level of both, semantics but also prosody. It should be compared whether effects of emotional content spoken with a neutral prosody are comparable to effects of emotional content with congruent prosody.

Another aspect that has to be mentioned is an unintended peculiarity of the neutral noun category: Most neutral words depict physical objects whereas positive and negative words are mostly describing concepts. Even though it is not obvious how this property of the neutral word category is effecting the observed neural differences, there is evidence that the concreteness of linguictic material is interacting with emotional processing (Kousta et al., 2009; Kousta et al., 2011; Vigliocco et al., 2014). It is, however, important to note that the selected words were matched in terms of imageability ratings, which speaks against a great difference in the abstractness of the emotion categories. Nevertheless, a systematic difference exist between the categories that deserve further attention in future studies.

Clearly, future research will be required to elucidate the reason for the differences across both studies and to investigate the boundary conditions of the N400-like effects to its full extent. It was hypothesized that the differences occurred due to context effects evoked by the unpredictability of the anticipated volume level of the next stimulus. Future research should further elucidate the boundary conditions of the occurring emotion effects in spoken word processing, e.g. under which conditions the AEC is preceded by the N400-like effect, in which contexts it is mainly driven by positive valence, if there is a modulation of the effect by arousal of the words or whether it does occur to adjectives and verbs as well. In a first step, it would be interesting to combine both studies of this dissertation into one experiment, being able to directly compare emotion effects in a condition with volume level manipulation to a condition without and thereby elucidating whether this was infact the causal agent for the additional N400-like effect.

In the present dissertation, first attempts were made to compare the neural pathways of the processing of emotional content in the written and spoken modality. Since EEG is a method that provides a very good temporal resolution but is restricted in its spatial resolution, it is not the most suitable method to detect the underlying neural mechanisms of the emotion effects. To further elucidate if emotion effects on word processing rely on a comparable neural system, combined EEG and fMRI would be a conceivable approach.

Future work should furthermore consider investigating more natural speech segments, for example as emotional words embedded into short neutral sentences. Single, isolated words are encountered rather seldom in natural settings and might be even more artificial for the spoken modality as compared to the visual modality, where we might encounter them in newspapers, advertisements etc. Nevertheless, for understanding emotion effects on a more complex level as in whole sentences or even longer speech segments, it is important to start at the very basis that is single spoken words.

In conclusion, in the present work I took decisive step in determining effects of emotional content modulating ERPs elicited by spoken word processing. With this dissertation, I filled the existing gap concerning research on spoken words with emotional content to a big extent. Clearly, future research is needed to finally clarify the boundary conditions and underlying neural mechanism of the revealed effects.

Appendix Table A

Seventy-two german nouns of three different emotion categories (positive, neutral and negative) selected as stimulus material (english translation in brackets) for written and auditory presentation

Positive Nouns	**Neutral Nouns**	**Negative Nouns**
Begegnung (encounter)	Batterie (battery)	Aas (carrion)
Ekstase (ecstasy)	Daumen (thumb)	Abfall (waste)
Erotik (eroticism)	Dokument (document)	Angeklagte (defendant)
Fee (fairy)	Eimer (bucket)	Bombe (bomb)
Geburt (birth)	Flasche (bottle)	Dieb (thief)
Geschenk (gift)	Hocker (stool)	Friedhof (cemetery)
Himbeere (rapsberry)	Karton (cardboard)	Gefängnis (prison)
Jubel (cheerings)	Kassette (cassette)	Grab (grave)
Kirsche (cherry)	Kiesel (pepple)	Granate (grenade)
Klavier (piano)	Kiste (box)	Käfig (cage)
Kunst (art)	Krawatte (tie)	Krankheit (disease)
Landschaft (landscape)	Linie (line)	Last (burden)
Lust (passion)	Magazin (magazine)	Leiche (corpse)
Ozean (ocean)	Pendel (pendulum)	Mord (murder)
Partner (partner)	Schrank (cupboard)	Munition (munitions)
Party (party)	Schraube (screw)	Sarg (casket)
Pracht (splendor)	Schwamm (sponge)	Schlacht (battle)
Reise (journey)	Stamm (trunk)	Schlag (stroke)
Schatz (treasure)	Strumpf (stocking)	Teufel (devil)
Sex (sex)	Stuhl (chair)	Torpedo (torpedo)
Spaß (fun)	Teppich (carpet)	Übelkeit (nausea)
Treffer (strike)	Tisch (table)	Überfall (attack)
Vorspiel (foreplay)	Wand (wall)	Verletzung (injury)
Wochenende (weekend)	Weste (vest)	Warze (wart)

6 References

Adler, G., Adler, J., 1991. Auditory stimulus processing at different stimulus intensities as reflected by auditory evoked potentials. BIOLOGICAL PSYCHIATRY, 29 (4), 347–356. doi:10.1016/0006-3223(91)90220-G.

Agrawal, D., Timm, L., Viola, F.C., Debener, S., Büchner, A., Dengler, R., Wittfoth, M., 2012. ERP evidence for the recognition of emotional prosody through simulated cochlear implant strategies. BMC Neurosci, 13, 113. doi:10.1186/1471-2202-13-113.

Alpers, G.W., Gerdes, Antje B M, Lagarie, B., Tabbert, K., Vaitl, D., Stark, R., 2009. Attention and amygdala activity: an fMRI study with spider pictures in spider phobia. Journal of neural transmission (Vienna, Austria : 1996), 116 (6), 747–757. doi:10.1007/s00702-008-0106-8.

Anders, S., Eippert, F., Weiskopf, N., Veit, R., 2008. The human amygdala is sensitive to the valence of pictures and sounds irrespective of arousal: an fMRI study. Social Cognitive and Affective Neuroscience, 3 (3), 233–243. doi:10.1093/scan/nsn017.

Balas, B., Conlin, C., 2015. The Visual N1 Is Sensitive to Deviations from Natural Texture Appearance. PloS one, 10 (9), e0136471. doi:10.1371/journal.pone.0136471.

Balconi, M., Pozzoli, U., 2005. Comprehending Semantic and Grammatical Violations in Italian. N400 and P600 Comparison with Visual and Auditory Stimuli. Journal of Psycholinguistic Research, 34 (1), 71–98. doi:10.1007/s10936-005-3633-6.

Bayer, M., Schacht, A., 2014. Event-related brain responses to emotional words, pictures, and faces - a cross-domain comparison. Frontiers in psychology, 5, 1106. doi:10.3389/fpsyg.2014.01106.

Bayer, M., Sommer, W., Schacht, A., 2010. Reading emotional words within sentences: the impact of arousal and valence on event-related potentials. International journal of psychophysiology : official journal of the International Organization of Psychophysiology, 78 (3), 299–307. doi:10.1016/j.ijpsycho.2010.09.004.

Bayer, M., Sommer, W., Schacht, A., 2011. Emotional words impact the mind but not the body: evidence from pupillary responses. Psychophysiology, 48 (11), 1554–1562. doi:10.1111/j.1469-8986.2011.01219.x.

Bayer, M., Sommer, W., Schacht, A., 2012a. Font size matters--emotion and attention in cortical responses to written words. PLoS ONE, 7 (5), e36042. doi:10.1371/journal.pone.0036042.

Bayer, M., Sommer, W., Schacht, A., 2012b. P1 and beyond: functional separation of multiple emotion effects in word recognition. Psychophysiology, 49 (7), 959–969. doi:10.1111/j.1469-8986.2012.01381.x.

Beagley, H.A., Knight, J.J., 1967. Changes in auditory evoked response with intensity. The Journal of laryngology and otology, 81 (8), 861–873. doi:10.1017/S0022215100067815.

Bedny, M., Thompson-Schill, S.L., 2006. Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. Brain and Language, 98 (2), 127–139. doi:10.1016/j.bandl.2006.04.008.

Begault, D.R., 1991. Preferred sound intensity increase for sensation of half distance. Perceptual and Motor Skills, 72, 1019–1029. doi:10.2466/pms.1991.72.3.1019.

Boersma, P., Weenik, D., 2009. Praat: doing phonetics by computer.

Bradley, M.M., Lang, P.J., 2000. Affective reactions to acoustic stimuli. Psychophysiology, 37 (2), 204–215. doi:10.1111/1469-8986.3720204.

Brandeis, D., Naylor, H., Halliday, R., Callaway, E., Yano, L., 1992. Scopolamine effects on visual information processing, attention and event-related potential map latencies. Psychophysiology, 29, 315–336.

Bröckelmann, A.-K., Steinberg, C., Elling, L., Zwanzger, P., Pantev, C., Junghöfer, M., 2011. Emotion-associated tones attract enhanced attention at early auditory processing: magnetoencephalographic correlates. The Journal of Neuroscience, 31 (21), 7801–7810. doi:10.1523/JNEUROSCI.6236-10.2011.

Cabeza, R., Nyberg, L., 2000. Imaging cognition II: An empirical review of 275 PET and fMRI studies. Journal of Cognitive Neuroscience, 12 (1), 1–47. doi:10.1162/08989290051137585.

Carrillo-de-la-Peña, M.T., 1999. Effects of intensity and order of stimuli presentation on AEPs: an analysis of the consistency of EP augmenting/reducing in the auditory modality. Clinical Neurophysiology, 110 (5), 924–932. doi:10.1016/S1388-2457(99)00041-3.

Chou, T.-L., Booth, J.R., Bitan, T., Burman, D.D., Bigio, J.D., Cone, N.E., Lu, D., Cao, F., 2006. Developmental and skill effects on the neural correlates of semantic processing to visually presented words. Human brain mapping, 27 (11), 915–924. doi:10.1002/hbm.20231.

Citron, F., Weekes, B., Ferstl, E., 2013. Effects of valence and arousal on written word recognition: time course and ERP correlates. Neuroscience Letters, 533, 90–95. doi:10.1016/j.neulet.2012.10.054.

Codispoti, M., DeCesarei, A., 2007. Arousal and attention: picture size and emotional reactions. Psychophysiology, 44 (5), 680–686. doi:10.1111/j.1469-8986.2007.00545.x.

Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J., 2000. Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. Biological psychology, 52 (2), 95–111. doi:10.1016/S0301-0511(99)00044-7.

Daltrozzo, J., Wioland, N., Kotchoubey, B., 2007. Sex differences in two event-related potentials components related to semantic priming. Archives of sexual behavior, 36 (4), 555–568. doi:10.1007/s10508-006-9161-0.

De Cesarei, A., Codispoti, M., 2006. When does size not matter? Effects of stimulus size on affective modulation. Psychophysiology, 43 (2), 207–215. doi:10.1111/j.1469-8986.2006.00392.x.

De Haan, A., Smit, M., Van der Stigchel, Stefan, Dijkerman, H.C., 2016. Approaching threat modulates visuotactile interactions in peripersonal space. Experimental brain research. doi:10.1007/s00221-016-4571-2.

Dehaene, S., Cohen, L., Sigman, M., Vinckier, F., 2005. The neural code for written words: a proposal. Trends in Cognitive Sciences, 9 (7), 335–341. doi:10.1016/j.tics.2005.05.004.

Diamond, E., Zhang, Y., 2016. Cortical processing of phonetic and emotional information in speech: A cross-modal priming study. Neuropsychologia, 82, 110–122. doi:10.1016/j.neuropsychologia.2016.01.019.

Erlbeck, H., Kübler, A., Kotchoubey, B., Veser, S., 2014. Task instructions modulate the attentional mode affecting the auditory MMN and the semantic N400. Frontiers in human neuroscience, 8, 654. doi:10.3389/fnhum.2014.00654.

Fecteau, S., Belin, P., Joanette, Y., Armony, J.L., 2007. Amygdala responses to nonlinguistic emotional vocalizations. NeuroImage, 36 (2), 480–487. doi:10.1016/j.neuroimage.2007.02.043.

Feng, W., Störmer, V.S., Martinez, A., McDonald, J.J., Hillyard, S.A., 2014. Sounds activate visual cortex and improve visual discrimination. The Journal of Neuroscience, 34 (29), 9817–9824. doi:10.1523/JNEUROSCI.4869-13.2014.

Fields, E.C., Kuperberg, G.R., 2012. It's All About You: an ERP study of emotion and self-relevance in discourse. NeuroImage, 62 (1), 562–574. doi:10.1016/j.neuroimage.2012.05.003.

Fischler, I., Bradley, M.M., 2006. Event-related potential studies of language and emotion: words, phrases, and task effects. Understanding Emotions, 156, 185–203. doi:10.1016/S0079-6123(06)56009-1.

Friederici, A.D., Alter, K., 2004. Lateralization of auditory language functions: A dynamic dual pathway model. Brain and Language, 89 (2), 267–276. doi:10.1016/S0093-934X(03)00351-1.

Friedman, L., Kenny, J.T., Wise, A.L., Wu, D., 1998. Brain Activation During Silent Word Generation Evaluated with Functional MRI. Brain and Language, 64 (2), 231–256.

Giraud, A.L., 2004. Contributions of Sensory Input, Auditory Search and Verbal Comprehension to Cortical Activity during Speech Processing. Cerebral Cortex, 14 (3), 247–255. doi:10.1093/cercor/bhg124.

Goydke, K.N., Altenmüller, E., Möller, J., Münte, T.F., 2004. Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. Brain research. Cognitive brain research, 21 (3), 351–359. doi:10.1016/j.cogbrainres.2004.06.009.

Hahne, A., Friederici, A.D., 2002. Differential task effects on semantic and syntactic processes as revealed by ERPs. Cognitive Brain Research, 13 (3), 339–356. doi:10.1016/S0926-6410(01)00127-6.

Hammond, J., 1990. Is learning to read and write the same as learning to speak? In F. Christie (Ed.), Literacy in a Changing World (pp. 26-53). Hawthorn, Victoria: ACE.

Hatzidaki, A., Baus, C., Costa, A., 2015. The way you say it, the way I feel it: emotional word processing in accented speech. Frontiers in psychology, 6, 351. doi:10.3389/fpsyg.2015.00351.

Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., Kissler, J., 2009. Amygdala activation during reading of emotional adjectives--an advantage for pleasant content. Social Cognitive and Affective Neuroscience, 4 (1), 35–49. doi:10.1093/scan/nsn027.

Herbert, C., Herbert, B.M., Ethofer, T., Pauli, P., 2011a. His or mine? The time course of self-other discrimination in emotion processing. Social neuroscience, 6 (3), 277–288. doi:10.1080/17470919.2010.523543.

Herbert, C., Herbert, B.M., Pauli, P., 2011b. Emotional self-reference: brain structures involved in the processing of words describing one's own emotions. Neuropsychologia, 49 (10), 2947–2956. doi:10.1016/j.neuropsychologia.2011.06.026.

Herbert, C., Junghofer, M., Kissler, J., 2008. Event related potentials to emotional adjectives during reading. Psychophysiology, 45 (3), 487–498. doi:10.1111/j.1469-8986.2007.00638.x.

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., Rockstroh, B., 2006. Processing of emotional adjectives: Evidence from startle EMG and ERPs. Psychophysiology, 43 (2), 197–206. doi:10.1111/j.1469-8986.2006.00385.x.

Herrmann, M.J., Huter, T., Plichta, M.M., Ehlis, A.-C., Alpers, G.W., Mühlberger, A., Fallgatter, A.J., 2008. Enhancement of activity of the primary visual cortex during processing of emotional stimuli as measured with event-related functional near-infrared spectroscopy and event-related potentials. Human brain mapping, 29 (1), 28–35. doi:10.1002/hbm.20368.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nature reviews. Neuroscience, 8 (5), 393–402. doi:10.1038/nrn2113.

Hillyard, S.A., Anllo-Vento, L., 1998. Event-related brain potentials in the study of visual selective attention. Proceedings of the National Academy of Sciences, 95 (3), 781–787. doi:10.1073/pnas.95.3.781.

Hinojosa, J.A., Mercado, F., Albert, J., Barjola, P., Peláez, I., Villalba-García, C., Carretié, L., 2015. Neural correlates of an early attentional capture by positive distractor words. Frontiers in psychology, 6, 24. doi:10.3389/fpsyg.2015.00024.

Hofmann, M.J., Kuchinke, L., Tamm, S., Võ, M.L.-H., Jacobs, A.M., 2009. Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. Cognitive, affective & behavioral neuroscience, 9 (4), 389–397. doi:10.3758/9.4.389.

Holmes, A., Nielsen, M.K., Green, S., 2008. Effects of anxiety on the processing of fearful and happy faces: an event-related potential study. Biological psychology, 77 (2), 159–173. doi:10.1016/j.biopsycho.2007.10.003.

Holt, D.J., Lynn, S.K., Kuperberg, G.R., 2009. Neurophysiological Correlates of Comprehending Emotional Meaning in Context. Journal of Cognitive Neuroscience, 21 (11), 2245–2262. doi:10.1162/jocn.2008.21151.

Ille, N., Berg, P., Scherg, M., 2002. Artifact Correction of the Ongoing EEG Using Spatial Filters Based on Artifact and Brain Signal Topographies. Journal of Clinical Neurophysiology, 19 (2), 113–124. doi:10.1097/00004691-200203000-00002.

Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malavade, K., Beattie, B., Leon, A.C., Stern, E., 1999. Linguistic threat activates the human amygdala. Proceedings of the National Academy of Sciences, 96 (18), 10456–10459. doi:10.1073/pnas.96.18.10456.

Jaspers-Fayer, F., Ertl, M., Leicht, G., Leupelt, A., Mulert, C., 2012. Single-trial EEG–fMRI coupling of the emotional auditory early posterior negativity. NeuroImage, 62 (3), 1807–1814. doi:10.1016/j.neuroimage.2012.05.018.

Johannes, S., Münte, T.F., Heinze, H.J., Mangun, G.R., 1995. Luminance and spatial attention effects on early visual processing. Cognitive Brain Research, 2 (3), 189–205. doi:10.1016/0926-6410(95)90008-X.

Johnstone, T., van Reekum, Carien M., Oakes, T.R., Davidson, R.J., 2006. The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. Social Cognitive and Affective Neuroscience, 1 (3), 242–249. doi:10.1093/scan/nsl027.

Junghöfer, M., Bradley, M.M., Elbert, T.R., Lang, P.J., 2001. Fleeting images: A new look at early emotion discrimination. Psychophysiology, 38, 175–178. doi:10.1111/1469-8986.3820175.

Junghöfer, M., Kissler, J., Schupp, H.T., Putsche, C., Elling, L., Dobel, C., 2010. A fast neural signature of motivated attention to consumer goods separates the sexes. Frontiers in human neuroscience, 4, 179. doi:10.3389/fnhum.2010.00179.

Junghöfer, M., Sabatinelli, D., Bradley, M.M., Schupp, H.T., Elbert, T.R., Lang, P.J., 2006. Fleeting images: rapid affect discrimination in the visual cortex. Neuroreport, 17 (2), 225–229. doi:10.1097/01.wnr.0000198437.59883.bb.

Kanske, P., Kotz, S.A., 2007. Concreteness in emotional words: ERP evidence from a hemifield study. Brain research, 1148, 138–148. doi:10.1016/j.brainres.2007.02.044.

Kasai, T., Morotomi, T., Katayama, J., Kumada, T., 2003. Attending to a location in three-dimensional space modulates early ERPs. Cognitive Brain Research, 17 (2), 273–285. doi:10.1016/S0926-6410(03)00115-0.

Keil, A., Moratti, S., Sabatinelli, D., Bradley, M.M., Lang, P.J., 2005. Additive effects of emotional content and spatial selective attention on electrocortical facilitation. Cerebral cortex (New York, N.Y. : 1991), 15 (8), 1187–1197. doi:10.1093/cercor/bhi001.

Kissler, J., Herbert, C., Peyk, P., Junghöfer, M., 2007. Buzzwords - Early cortical responses to emotional words during reading, 18 (6), 475–480. doi:10.1111/j.1467-9280.2007.01924.x.

Kissler, J., Herbert, C., Winkler, I., Junghofer, M., 2009. Emotion and attention in visual word processing: an ERP study. Biological psychology, 80 (1), 75–83. doi:10.1016/j.biopsycho.2008.03.004.

Kotz, S.A., Paulmann, S., 2007. When emotional prosody and semantics dance cheek to cheek: ERP evidence. Brain research, 1151, 107–118. doi:10.1016/j.brainres.2007.03.015.

Kousta, S.-T., Vigliocco, G., Vinson, D.P., Andrews, M., Del Campo, E., 2011. The representation of abstract words: why emotion matters. Journal of experimental psychology. General, 140 (1), 14–34. doi:10.1037/a0021446.

Kousta, S.-T., Vinson, D.P., Vigliocco, G., 2009. Emotion words, regardless of polarity, have a processing advantage over neutral words. Cognition, 112 (3), 473–481. doi:10.1016/j.cognition.2009.06.007.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annual review of psychology, 62, 621–647. doi:10.1146/annurev.psych.093008.131123.

Lane, R.D., Chua, P.M.-L., Dolan, R.J., 1999. Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. Neuropsychologia, 37 (9), 989–997. doi:10.1016/S0028-3932(99)00017-2.

Lang, P., 1995. The emotion probe: Studies of motivation and attention. AMERICAN PSYCHOLOGIST, 50 (5), 372–385. doi:10.1037//0003-066X.50.5.372.

Lang, P.J., Bradley, M.M., 2010. Emotion and the motivational brain. Biological psychology, 84 (3), 437–450. doi:10.1016/j.biopsycho.2009.10.007.

Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1997. Motivated attention: Affect, activation, and action. In P. J. Lang, R. F. Simon, & R. T. Balaban (Eds.), Attention and orienting: Sensory and motivational, Attention and orienting: Sensory and motivational.

Lang, P.J., Bradley, M.M., Fitzsimmons, J.R., Cuthbert, B.N., Scott, J.D., Moulder, B., Nangia, V., 1998. Emotional arousal and activation of the visual cortex: An fMRI analysis. Psychophysiology, 35 (2), 199–210. doi:10.1017/S0048577298001991.

Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalography and Clinical Neurophyiology, 48, 609–621.

Liu, B., Jin, Z., Wang, Z., Hu, Y., 2010. The interaction between pictures and words: evidence from positivity offset and negativity bias. Experimental brain research, 201 (2), 141–153. doi:10.1007/s00221-009-2018-8.

Lück, M., Hahne, A., Clahsen, H., 2006. Brain potentials to morphologically complex words during listening. Brain research, 1077 (1), 144–152. doi:10.1016/j.brainres.2006.01.030.

Marcus, S.M., Frauenfelder, U., 2007. Word recognition - uniqueness or deviation? A theoretical note. Language and Cognitive Processes, 1 (2), 163–169. doi:10.1080/01690968508402077.

Marinkovic, K., Dhond, R.P., Dale, A.M., Glessner, M., Carr, V., Halgren, E., 2003. Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing. Neuron, 38 (3), 487–497. doi:10.1016/S0896-6273(03)00197-1.

Markowitsch, H.J., 1998. Differential contribution of right and left amygdala to affective information processing. Behavioural Neurology, 11 (4), 233–244.

Marslen-Wilson, W.D., Tyler, L.K., 1980. The temporal structure of spoken language understanding. Cognition, 8 (1), 1–71.

Marslen-Wilson, W.D., Welsh, A., 1978. Processing Interactions and Lexical Access during Word Recognition in Continuous Speech. Cognitive Psychology, 10 (1), 29–63. doi:10.1016/0010-0285(78)90018-X.

Martin, R.C., 2003. Language processing: functional organization and neuroanatomical basis. Annual review of psychology, 54, 55–89. doi:10.1146/annurev.psych.54.101601.145201.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping

(ICBM). Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 356 (1412), 1293–1322. doi:10.1098/rstb.2001.0915.

McClelland, J.L., Rumelhart, D.E., 1981. An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. Psychological Review, 88 (5), 375–407. doi:10.1037/0033-295X.88.5.375.

McDermott, K.B., Petersen, S.E., Watson, J.M., Ojemann, J.G., 2003. A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. Neuropsychologia, 41 (3), 293–303. doi:10.1016/S0028-3932(02)00162-8.

McDonald, J.J., Störmer, V.S., Martinez, A., Feng, W., Hillyard, S.A., 2013. Salient sounds activate human visual cortex automatically. The Journal of Neuroscience, 33 (21), 9194–9201. doi:10.1523/JNEUROSCI.5902-12.2013.

Mesulam, M., 1998. From sensation to cognition. BRAIN, 121 (6), 1013–1052. doi:10.1093/brain/121.6.1013.

Metzger, W., 1942. Das Räumliche der Hör- und Sehwelt bei der Rundfunkübertragung, Berlin: Schenk.

Mitchell, D., Luo, Q., Mondillo, K., Vythilingam, M., Finger, E.C., Blair, R. J. R., 2008. The interference of operant task performance by emotional The interference of operant task performance by emotional distracters: An antagonistic relationship between the amygdala and frontoparietal cortices. NeuroImage, 40 (2), 859–868. doi:10.1016/j.neuroimage.2007.08.002.

Mittermeier, V., Leicht, G., Karch, S., Hegerl, U., Möller, H.-J., Pogarell, O., Mulert, C., 2011. Attention to emotion: auditory-evoked potentials in an emotional choice reaction task and personality traits as assessed by the NEO FFI. European archives of psychiatry and clinical neuroscience, 261 (2), 111–120. doi:10.1007/s00406-010-0127-9.

Nakayama, M., Fujimoto, M., 2015. Features of Oculo-motors and their chronological changes in response to varying sizes of visual stimuli. Multimed Tools Appl, 74 (8), 2841–2859. doi:10.1007/s11042-013-1824-y.

Nakic, M., Smith, B.W., Busis, S., Vythilingam, M., Blair, R.J.R., 2006. The impact of affect and frequency on lexical decision: the role of the amygdala and inferior frontal cortex. NeuroImage, 31 (4), 1752–1761. doi:10.1016/j.neuroimage.2006.02.022.

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Human brain mapping, 15 (1), 1–25. doi:10.1002/hbm.1058.

Ofek, E., Purdy, S.C., Ali, G., Webster, T., Gharahdaghi, N., McCann, C.M., 2013. Processing of emotional words after stroke: an electrophysiological study. Clinical neurophysiology :

official journal of the International Federation of Clinical Neurophysiology, 124 (9), 1771–1778. doi:10.1016/j.clinph.2013.03.005.

Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia, 9, 97–113.

Opitz, B., Degner, J., 2012. Emotionality in a second language: it's a matter of time. Neuropsychologia, 50 (8), 1961–1967. doi:10.1016/j.neuropsychologia.2012.04.021.

Palazova, M., Mantwill, K., Sommer, W., Schacht, A., 2011. Are effects of emotion in single words non-lexical? Evidence from event-related brain potentials. Neuropsychologia, 49 (9), 2766–2775. doi:10.1016/j.neuropsychologia.2011.06.005.

Palazova, M., Sommer, W., Schacht, A., 2013. Interplay of emotional valence and concreteness in word processing: An event-related potential study with verbs. Brain and Language, 125 (3), 264–271. doi:10.1016/j.bandl.2013.02.008.

Park, Y.-M., Lee, S.-H., Kim, S., Bae, S.-M., 2010. The loudness dependence of the auditory evoked potential (LDAEP) in schizophrenia, bipolar disorder, major depressive disorder, anxiety disorder, and healthy controls. Progress in neuro-psychopharmacology & biological psychiatry, 34 (2), 313–316. doi:10.1016/j.pnpbp.2009.12.004.

Partala, T., Surakka, V., 2003. Pupil size variation as an indication of affective processing. International Journal of Human-Computer Studies, 59 (1-2), 185–198. doi:10.1016/S1071-5819(03)00017-X.

Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Meth. Find. Exp. Clin. Pharmacol. 24, 5–12. Methods & Findings in Experimental & Clinical Pharmacology, 24, 5–12.

Paulmann, S., Bleichner, M., Kotz, S.A., 2013. Valence, arousal, and task effects in emotional prosody processing. Front Psychol, 4, 345. doi:10.3389/fpsyg.2013.00345.

Paulmann, S., Kotz, S.A., 2008. Early emotional prosody perception based on different speaker voices. Neuroreport, 19 (2), 209–213. doi:10.1097/WNR.0b013e3282f454db.

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. doi:10.1016/j.brainres.2009.07.102.

Paulus, A., Wentura, D., 2014. Threatening joy: approach and avoidance reactions to emotions are influenced by the group membership of the expresser. Cognition & emotion, 28 (4), 656–677. doi:10.1080/02699931.2013.849659.

Paulus, A., Wentura, D., 2016. It depends: Approach and avoidance reactions to emotional expressions are influenced by the contrast emotions presented in the task. Journal of experimental psychology. Human perception and performance, 42 (2), 197–212. doi:10.1037/xhp0000130.

Perrin, F., García-Larrea, L., 2003. Modulation of the N400 potential during auditory phonological/semantic interaction. Cognitive Brain Research, 17 (1), 36–47. doi:10.1016/S0926-6410(03)00078-8.

Pfabigan, D.M., Sailer, U., Lamm, C., 2015. Size does matter! Perceptual stimulus properties affect event-related potentials during feedback processing. Psychophysiology, 52 (9), 1238–1247. doi:10.1111/psyp.12458.

Picton, T.W., Goodman, W.S., Bryce, D.P., 1970. Amplitude of evoked responses to tones of high intensity. ACTA OTO-LARYNGOLOGICA, 70 (2), 77–82.

Pinheiro, A.P., Del Re, E., Mezin, J., Nestor, P.G., Rauber, A., McCarley, R.W., Gonçalves, O.F., Niznikiewicz, M.A., 2013. Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. Psychological medicine, 43 (3), 603–618. doi:10.1017/S003329171200133X.

Plichta, M.M., Gerdes, A B M, Alpers, G.W., Harnisch, W., Brill, S., Wieser, M.J., Fallgatter, A.J., 2011. Auditory cortex activation is modulated by emotion: a functional near-infrared spectroscopy (fNIRS) study. NeuroImage, 55 (3), 1200–1207. doi:10.1016/j.neuroimage.2011.01.011.

Poirier, C., Collignon, O., Devolder, A.G., Renier, L., Vanlierde, A., Tranduy, D., Scheiber, C., 2005. Specific activation of the V5 brain area by auditory motion processing: an fMRI study. Brain research. Cognitive brain research, 25 (3), 650–658. doi:10.1016/j.cogbrainres.2005.08.015.

Pourtois, G., Grandjean, D., Sander, D., Vuilleumier, P., 2004. Electrophysiological correlates of rapid spatial orienting towards fearful faces. Cerebral cortex (New York, N.Y. : 1991), 14 (6), 619–633. doi:10.1093/cercor/bhh023.

Pourtois, G., Schettino, A., Vuilleumier, P., 2013. Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. Biological psychology, 92 (3), 492–512. doi:10.1016/j.biopsycho.2012.02.007.

Price, C.J., Wise, R., Watson, J., Patterson, K., Howard, D., Frackowiak, R., 1994. Brain activity during reading - The effects of exposure duration and task. BRAIN, 117 (6), 1255–1269. doi:10.1093/brain/117.6.1255.

Pulvermüller, F., Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Progress in neurobiology, 79 (1), 49–71. doi:10.1016/j.pneurobio.2006.04.004.

Qin, W., Yu, C., 2013. Neural pathways conveying novisual information to the visual cortex. Neural plasticity, 2013, 864920. doi:10.1155/2013/864920.

Rapin, I., Schimmel, H., Tourk, L.M., Krasnegor, N.A., Pollak, C., 1966. Evoked responses to clicks and tones of varying intensity in waking adults. Electroencephalography and Clinical Neurophyiology, 21 (4), 335–344. doi:10.1016/0013-4694(66)90039-3.

Recio, G., Sommer, W., Schacht, A., 2011. Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. Brain research, 1376, 66–75. doi:10.1016/j.brainres.2010.12.041.

Relander, K., Rama, P., Kujala, T., 2009. Word Semantics Is Processed Even without Attentional Effort. Journal of Cognitive Neuroscience, 21 (8), 1511–1522. doi:10.1162/jocn.2009.21127.

Rellecke, J., Palazova, M., Sommer, W., Schacht, A., 2011. On the automaticity of emotion processing in words and faces: event-related brain potentials evidence from a superficial task. Brain and cognition, 77 (1), 23–32. doi:10.1016/j.bandc.2011.07.001.

Rellecke, J., Sommer, W., Schacht, A., 2012. Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. Biological psychology, 90 (1), 23–32. doi:10.1016/j.biopsycho.2012.02.002.

Rohr, L., Abdel Rahman, R., 2015. Affective responses to emotional words are boosted in communicative situations. NeuroImage, 109, 273–282. doi:10.1016/j.neuroimage.2015.01.031.

Sabatinelli, D., Lang, P.J., Bradley, M.M., Costa, V.D., Keil, A., 2009. The timing of emotional discrimination in human amygdala and ventral visual cortex. The Journal of Neuroscience, 29 (47), 14864–14868. doi:10.1523/JNEUROSCI.3278-09.2009.

Sakai, K.L., 2005. Language acquisition and brain development. Science, 310 (5749), 815–819. doi:10.1126/science.1113530.

Sauter, D.A., Eimer, M., 2010. Rapid Detection of Emotion from Human Vocalizations. Journal of Cognitive Neuroscience, 22 (3), 474–481. doi:10.1162/jocn.2009.21215.

Schacht, A., Sommer, W., 2009a. Emotions in word and face processing: early and late cortical responses. Brain and cognition, 69 (3), 538–550. doi:10.1016/j.bandc.2008.11.005.

Schacht, A., Sommer, W., 2009b. Time course and task dependence of emotion effects in word processing. Cognitive, affective & behavioral neuroscience, 9 (1), 28–43. doi:10.3758/CABN.9.1.28.

Schadow, J., Lenz, D., Thaerig, S., Busch, N.A., Fründ, I., Herrmann, C.S., 2007. Stimulus intensity affects early sensory processing: sound intensity modulates auditory evoked gamma-band activity in human EEG. International journal of psychophysiology : official journal of the International Organization of Psychophysiology, 65 (2), 152–161. doi:10.1016/j.ijpsycho.2007.04.006.

Scherer, K., 2003. Vocal communication of emotion: A review of research paradigms. Speech Communication, 40 (1-2), 227–256. doi:10.1016/S0167-6393(02)00084-5.

Schettino, A., Keil, A., Porcu, E., Müller, M.M., 2016. Shedding light on emotional perception: Interaction of brightness and semantic content in extrastriate visual cortex. NeuroImage, 133, 341–353. doi:10.1016/j.neuroimage.2016.03.020.

Schindler, S., Kissler, J., 2016. Selective visual attention to emotional words: Early parallel frontal and visual activations followed by interactive effects in visual cortex. Human brain mapping, 37 (10), 3575–3587. doi:10.1002/hbm.23261.

Schindler, S., Wegrzyn, M., Steppacher, I., Kissler, J., 2015. Perceived communicative context and emotional content amplify visual word processing in the fusiform gyrus. The Journal of Neuroscience, 35 (15), 6010–6019. doi:10.1523/JNEUROSCI.3346-14.2015.

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. doi:10.3758/s13415-012-0132-8.

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. doi:10.1162/089892903322598102.

Schupp, H.T., Junghöfer, M., Weike, A.I., Hamm, A.O., 2003. Emotional Facilitation of Sensory Processing in the Visual Cortex. Psychological Science, 14 (1), 7–13. doi:10.1111/1467-9280.01411.

Schupp, H.T., Junghöfer, M., Weike, A.I., Hamm, A.O., 2004. The selective processing of briefly presented affective pictures: an ERP analysis. Psychophysiology, 41 (3), 441–449. doi:10.1111/j.1469-8986.2004.00174.x.

Schupp, H.T., Schmälzle, R., Flaisch, T., Weike, A.I., Hamm, A.O., 2012. Affective picture processing as a function of preceding picture valence: an ERP analysis. Biological psychology, 91 (1), 81–87. doi:10.1016/j.biopsycho.2012.04.006.

Schupp, H.T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A.I., Hamm, A.O., 2007. Selective visual attention to emotion. J. Neurosci., 27 (5), 1082–1089. doi:10.1523/JNEUROSCI.3223-06.2007.

Schwartz, B.L., Krantz, J.H., 2016. Sensation and Perception, Los Angeles: Sage.

Schwartz, R., Pell, M.D., 2012. Emotional Speech Processing at the Intersection of Prosody and Semantics. PloS one, 7 (10). doi:10.1371/journal.pone.0047279.

Scott, G.G., O'Donnell, P.J., Leuthold, H., Sereno, S.C., 2009. Early emotion word processing: evidence from event-related potentials. Biological psychology, 80 (1), 95–104. doi:10.1016/j.biopsycho.2008.03.010.

Scott, S.K., Blank, C.C., Rosen, S., Wise, R., 2000. Identification of a pathway for intelligible speech in the left temporal lobe. BRAIN, 123 (12), 2400–2406. doi:10.1093/brain/123.12.2400.

Seghier, M.L., Lazeyras, F., Pegna, A.J., Annoni, J.-M., Zimine, I., Mayer, E., Michel, C.M., Khateb, A., 2004. Variability of fMRI activation during a phonological and semantic language task in healthy subjects. Human brain mapping, 23 (3), 140–155. doi:10.1002/hbm.20053.

Sekihara, K., Sahani, M., Nagarajan, S.S., 2005. Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. NeuroImage, 25 (4), 1056–1067. doi:10.1016/j.neuroimage.2004.11.051.

Stein, J., 2001. The magnocellular theory of developmental dyslexia. Dyslexia, 7 (1), 12–36. doi:10.1002/dys.186.

Steinhauer, K., Alter, K., Friederici, A.D., 1999. Brain potentials indicate immediate Brain potentials indicate immediate use of prosodic cues in natural speech processing. Nature Neuroscience, 2 (2), 191–196. doi:10.1038/5757.

Tabert, M.H., Borod, J.C., Tang, C.Y., Lange, G., Wei, T.C., Johnson, R., Nusbaum, A.O., Buchsbaum, M.S., 2001. Differential amygdala activation during emotional decision and recognition memory tasks using unpleasant words: an fMRI study. Neuropsychologia, 39 (6), 556–573. doi:10.1016/S0028-3932(00)00157-3.

Thaerig, S., Behne, N., Schadow, J., Lenz, D., Scheich, H., Brechmann, Andre, Herrmann, Christoph S., 2008. Sound level dependence of auditory evoked potentials: Simultaneous EEG recording and low-noise fMRI. International Journal of psychophysiology, 67 (3), 235–241. doi:10.1016/j.ijpsycho.2007.06.007.

Thierry, G., Roberts, M.V., 2007. Event-related potential study of attention capture by affective sounds. Neuroreport, 18 (3), 245–248. doi:10.1097/WNR.0b013e328011dc95.

Valdés-Conroy, B., Román, F.J., Hinojosa, J.A., Shorkey, S.P., 2012. So far so good: emotion in the peripersonal/extrapersonal space. PloS one, 7 (11), e49162. doi:10.1371/journal.pone.0049162.

Valdés-Conroy, B., Sebastián, M., Hinojosa, J.A., Román, F.J., Santaniello, G., 2014. A close look into the near/far space division: a real-distance ERP study. Neuropsychologia, 59, 27–34. doi:10.1016/j.neuropsychologia.2014.04.009.

Vigliocco, G., Kousta, S.-T., Della Rosa, P.A., Vinson, D.P., Tettamanti, M., Devlin, J.T., Cappa, S.F., 2014. The neural representation of abstract words: the role of emotion. Cerebral cortex (New York, N.Y. : 1991), 24 (7), 1767–1777. doi:10.1093/cercor/bht025.

Võ, M.L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M.J., Jacobs, A.M., 2009. The Berlin Affective Word List Reloaded (BAWL-R). Behavior research methods, 41 (2), 534–538. doi:10.3758/BRM.41.2.534.

VonBekesy, G., 1949. The Moon Illusion and Similar Auditory Phenomena. American Journal of Psychology, 62 (4), 540–552. doi:10.2307/1418558.

Vorobyev, V.A., Alho, K., Medvedev, S.V., Pakhomov, S.V., Roudas, M.S., Rutkovskaya, J.M., Tervaniemi, M., Van Zuijen, Titia L, Näätänen, R., 2004. Linguistic processing in visual and modality-nonspecific brain areas: PET recordings during selective attention. Brain research. Cognitive brain research, 20 (2), 309–322. doi:10.1016/j.cogbrainres.2004.03.011.

Vuilleumier, P., Driver, J., 2007. Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 362 (1481), 837–855. doi:10.1098/rstb.2007.2092.

Vuilleumier, P., Richardson, M.P., Armony, J.L., Driver, J., Dolan, R.J., 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nature Neuroscience, 7 (11), 1271–1278. doi:10.1038/nn1341.

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. doi:10.1016/j.cogbrainres.2004.03.015.

Wicha, N.Y., Bates, E.A., Moreno, E.M., Kutas, M., 2003. Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. Neuroscience Letters, 346 (3), 165–168. doi:10.1016/S0304-3940(03)00599-8.

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. doi:10.1016/j.neuroimage.2007.09.028.

Wilson, S.M., Saygin, A.P., Sereno, M.I., Iacoboni, M., 2004. Listening to speech activates motor areas involved in speech production. Nature Neuroscience, 7 (7), 701–702. doi:10.1038/nn1263.

Wirth, M., Horn, H., Koenig, T., Stein, M., Federspiel, A., Meier, B., Michel, C.M., Strik, W., 2007. Sex differences in semantic processing: event-related brain potentials distinguish between lower and higher order semantic analysis during word reading. Cerebral cortex (New York, N.Y. : 1991), 17 (9), 1987–1997. doi:10.1093/cercor/bhl121.

Zahorik, P., 1996. Auditory Distance Perception: A Literature Review, Wisconsin: University of Wisconsin-Madison.

Zimmer, U., Lewald, J., Erb, M., Grodd, W., Karnath, H.-O., 2004. Is there a role of visual cortex in spatial hearing? The European journal of neuroscience, 20 (11), 3148–3156. doi:10.1111/j.1460-9568.2004.03766.x.

7 Acknowledgments

First of all, my gratitude goes to Annekathrin Schacht, who made this thesis possible and who accompanied me on this way. You encouraged me to work independently on my research question and in the same way supported me and enlightened my thoughts whenever necessary. Thank you for given me the opportunity and supporting me in asserting myself in a field outside my original subject. Your enthusiasm impressed me from the first day and cheered me up in times of doubts. Furthermore, I would like to thank Igor Kagan and Julia Fischer for co-supervising my thesis, for being discussion partners and asking critical questions during our meetings.

I want to extent my gratitude to my whole research group. The varying group composition in the last years enlarges the list of people I have to thank but in the same way have enlarged my list of friends in (and outside) of Göttingen. First of all, I would like to thank Wiebke, for being the best office neighbor, a wonderful conference travel company and for assisting me in distress whenever needed. I am grateful that you became a great friend to me throughout the years. I would particularly thank Mareike for both her personal and professional support, for all her valuable suggestions and for discussing data (and other important topics) in the evening on the balcony. I would like to thank Adi, for being "The Admiral" sometimes and reminding us that we have to relax (at least ones a week). Special thanks go to Anna (for being good-humored all the time and fortunately never leaving Göttingen and our group completely), Rebecca (for her assistance in bioacoustical analyses and all other important questions), Katja, Svenja, Julia, Rebecca, Lena and all other group members I cannot list here, who ensured a warm-hearted and humorous atmosphere throughout the whole time.

There are a lot of other people (too many to mention here), who supported me during the last years, who encouraged or distracted me (whatever was necessary at that specific moment). Thank you all! Special thanks go to my "Hüpfburgen-Business" for reminding me there are other, major important things in life outside of academia. I would like to express my special thanks to Amelie, the best friend I could imagine and to Katja & Kirsten, for accompanying me the whole duration of study and for all the skype calls showing me, I am not alone with all this. Last but not least, my biggest gratitude goes to my family for their continual support and their unfailing trust in me. Without knowing you being there for me and backing me up, all this would not have been possible. DANKE!

Annika Graß

EDUCATION

10/2013 – 10/2016 University of Göttingen, Behavior and Cognition PhD program

PhD in Psychology

Project: The Influence of Emotional Content on Event-Related Brain Potentials during Spoken Word Processing

10/2011-05/2013 International Max Planck Research School, University of Göttingen

MSc/PhD/MD-PhD Neurosciences Program (Grade: B -1.7)

Thesis: Effects of Emotional Meaning in Written and Spoken Word Processing - An Event-Related Potential Study

10/2008-10/2011 University of Göttingen,

B.Sc., Biology - First Class Honors (Grade: 1.5)

Thesis: The effect of professionalism on recognition of acted emotion in speech

08/2001-07/2008 Georg-Büchner-Gymnasium, Letter

Abitur (Grade: 2.4)

TEACHING AND ADMINISTRATIVE EXPERIENCE

08/2016	Teacher of the practical course “Biocognition”
04/2014-05/2015	Program Committee of “Behavior and Cognition” PhD Program as Student representative
10/2013-02/2014	Teaching assistant for the lecture “Good scientific practice”, conceptualizing the student’s exam, revising the exam
2013-2014	Supervision of two master theses

INTERNSHIPS / LAB ROTATIONS

09/01/2012 – 26/02/2012 Labrotation at Prof. J. Frahm, Biomedical NMR, Max Planck Institute for Biophysical Chemistry, Göttingen

Topic: „Exploring the attentional influence on somatosensation using fMRI“

27/02/2012 – 27/04/2012	Labrotation at Prof. Stefan Treue and Dr. Igor Kagan, Decision and Awareness Group, German Primate Center, Göttingen Topic: „Effort-based Decision Making in Non-human Primates”
30/04/2012 – 22/06/2012	Labrotation at Prof. Paulus, Clinical Neurophysiology, Medical Faculty of the University of Göttingen, Göttingen Topic: „The effect of transcranial Laser stimulation on cortical excitability“

PUBLICATIONS

Grass, A.; Bayer, M.; Schacht, A. (2016) - Electrophysiological Correlates of Emotional Content and Volume Level in Spoken Word Processing. Frontiers in Human Neuroscience 19, 347-357. doi: 10.3389/fnhum.2016.00326

Grass, A.; Bayer, M.; Rossi, V.; Hammerschmidt, W.; Schacht, A. (in preparation) - Effects of emotional content in written and spoken word processing: Evidence from event-related brain potentials.

Rossi, V; Vanlessen, N; Bayer, M; **Grass, A**; Pourtois, G; Schacht, A. (in press) - Motivational salience modulates early visual cortex responses across task sets. Journal of Cognitive Neuroscience. doi: 10.1162/jocn_a_01093

Bayer, M; Rossi, V; Vanlessen, N; **Grass, A**; Schacht, A; Pourtois, G. (2016) - Independent effects of motivation and spatial attention in the human visual cortex. Social Cognitive and Affective Neuroscience. doi: https://doi.org/10.1093/scan/nsw162

Juergens,R.; **Grass, A**.; Drolet, M.; Fischer, J. (2015) - Effect of Acting Experience on Emotion Expression and Recognition in Voice: Non-Actors Provide Better Stimuli than Expected. Journal of Nonverbal Behavior 39 (3), 195 -214. doi:10.1007/s10919-015-0209-5

POSTER PRESENTATIONS

Grass, A.; Bayer, M.; Schacht, A. – „Effects of volume level and emotional content on spoken word processing” (3rd International Conference European Society for Cognition and Affective Neuroscience; Porto, Portugal, 23th – 26th of June, 2016)

Grass, A.; Bayer, M.; Rossi, V.; Hammerschmidt, W.; Schacht, A. - “Effects of emotional content in written and spoken word processing: Evidence from event-related brain potentials” (Annual meeting of the cognitive Neuroscience Society 2016, New York, USA, 2nd – 5th of April, 2016)

Schacht, A.; **Grass, A.**; Grimm, A.; Bayer, M. – Emodulation of visual sensory processing by associated valence – Evidence from event-related brain potentials“ (Annual meeting of the cognitive Neuroscience Society 2016, New York, USA, 2nd – 5th of April, 2016)

Rossi, V.; Vanlessen, N.; Bayer, M.; Grass, A.; Schacht, A.; Pourtois, G. – „Incentive motivation alters early sensory processing in visual cortex: Evidence from ERPs“(Annual meeting of the cognitive Neuroscience Society 2016, New York, USA, 2nd – 5th of April, 2016)

Grass, A.; Hammerschmidt, W.; Bayer, M.; Schacht, A. - "Effects of emotional content in written and spoken word processing: Evidence from event-related brain potentials" (Summerschool "Emotion Expressions in Human and Non-Human Communication", Göttingen, Germany, 21st – 25th of September, 2015)

Grass, A.; Hammerschmidt, W.; Bayer, M.; Schacht, A. - "Cross-modal comparison of event-related brain potentials to emotional meaning in written and spoken language processing" (Neuronus 2015, IBRO & IRUN Neuroscience Forum; Krakow, Poland, 17th – 19th of April, 2015)

Grass, A.; Bayer, M.; Vanlessen, N.; Rossi, V.; Pourtois, G.; Schacht, A. – "The influence of reward motivation and spatial attention on early visual processing - Behavioural and pupillometric data" (U4 Workshop Extra-retinal influences on vision: Vistas, methods, and current controversies"; Göttingen, Germany, 18th – 19th of March, 2015)

Bayer, M.; **Grass, A**.; Rossi, V.; Vanlessen, N.; Pourtois, G.; Schacht, A. – "The influence of reward motivation and spatial attention on early visual processing – Event-related potentials" (U4 Workshop Extra-retinal influences on vision: Vistas, methods, and current controversies"; Göttingen, Germany, 18th – 19th of March, 2015)

Grass, A.; Hammerschmidt, W.; Bayer, M.; Schacht, A. - "Effects of emotional meaning in language processing: Evidence from event-related brain potentials in cross-modal comparison" (57th Conference of Experimental Psychologists; Hildesheim, Germany, 8th – 11th of March, 2015)

Grass, A.; Hammerschmidt, W.; Bayer, M.; Schacht, A. - "Emotion effects in spoken and written word processing: Evidence from ERPs in cross-modal comparison" (2nd International Conference European Society for Cognition and Affective Neuroscience; Dortmund, Germany, 7th – 10th of May, 2014)

TALKS

Grass, A. – "Effects of emotional content in written and spoken word processing: Evidence from event-related brain potentials" (U4 Workshop "Emotional, attentional and motivational influences on visual perception", Ghent, Belgium, 26th-27th May, 2016