

Mapping Speech Sound to Mental Representation:
Neurophonological Evidence from
Event-Related Brain Potentials

Dissertation
zur Erlangung des akademischen Grades
des Doktors der Naturwissenschaften
an der Universität Konstanz
im Fachbereich Psychologie

vorgelegt von
Dipl. Psych. Sonia Cornell

Tag der mündlichen Prüfung: 30. April 2010

Referent: Prof. C. Eulitz

Referentin: Prof. A. Lahiri

Thank you (Acknowledgements)....

I am grateful to both my supervisors, **Prof. Dr. Carsten Eulitz** and **Prof. Dr. Aditi Lahiri** for their involvement in developing the 'big picture'. Towards that big picture, they contributed to concepts of experimental designs, ideas, and theoretical linguistic-neuroscientific backgrounds.

In particular, I would like to thank **Carsten Eulitz** for his many hours of patience, the valuable theoretical, technical, and person-to-person discussions, and especially being able to learn from him.

I want to thank **Aditi Lahiri** for believing in me and expanding my self-confidence with her warm-hearted support. I deeply felt the loss after she departed Constance but am confident that I will continue to learn from her.

But my appreciation extends even farther to:

My colleagues and predecessors **Mathias Scharinger**, **Verena Felder**, and **Frank Zimmerer** for their neighborly presence, friendship, and all the named and un-named help they gave in the past three years.

My colleague **Alex Bobrov**, without whom I never could have conducted my experiments so quickly, I am grateful for his incredible fast programming-arts, his fast 'gelling' during the EEG preparation, and his dedicated engagement. My research assistants, **Gabriela Salagan**, **Natalia Bekemeier**, and **Ramona Grutschnig**, who have done so much of the necessary and unrecognized research for this project.

Verena Felder, **Mathias Scharinger**, **Frank Zimmerer**, **Allison Wetterlin**, **Andrea Myers** and **Tanja Rinker** for their comments, proof readings, and feedback which are so critical for any research effort.

My colleagues **KaiRobin Grzyb**, **Martin Meier**, **Ronny Hannemann**, **Eva Smolka**, **Nadine Tema**, and **Marco Steinhauser** for their support and valuable discussions, ideas, statistics, literature, images, and their open doors and ears. **Patric Berg**, who showed me with a calm and patient way of being, so that I could manage computer programs such as BESA and that behind every problem there is a solution. **Henning Reetz** for his help with my speech stimuli, his scripts for determining the acoustic characteristics as well as the frequency values of my (squeaks and) sounds. **Christian Scharinger** for the second experiment! For those many answers concerning the experiment, 'when do you start with your dissertation?'. **Bertram Opitz**, without whose encouragement I would never have come to Constance, and whose feedback, discussions and brain nurturing have left an indelible mark.

My dear friends **Katrin Lazar** and **Björn Graf**, without whom Constance would not have been Constance and whose house was my house and whose kitchen was my kitchen. Thank you for the sisterhood! **Juliane Hofman** for her wonderful friendship and accompaniment through the exhausting time, whose belief in me made me continue....the many Skype hours...thank you for your friendship! **Anne Schulze** for her vibrancy and the entrance into the Soul of Constance, for her acquaintance and enterprise. **Anni** and **Dieter Kallinke** for staying tuned into me and stuck with me.

For the continuing and unquestioning support I will always be grateful to my **family**, **my mom** and **dad**, **brother & sister-in-law**, my one-and-only **grandmother**, and my **Cologne-family**!

Table of Contents

Glossary (Abbreviation Index)	V
Figure & Table Legend	VI
0 Abstract	IX
0.1 Deutsche Zusammenfassung (German Summary)	XIII
1 General Introduction	1
1.1 Process of Word Recognition	1
1.2 Models of Speech Perception	8
1.2.1 Usage-based Model	8
1.2.2 Shortlist	11
1.2.3 Distributed Model of Speech Perception	14
1.2.4 Featurally Underspecified Lexicon Model	19
1.3 Underspecified Representations	26
1.3.1 Evidence from Psycholinguistic Behavioral Experiments	26
1.3.2 Electrophysiological Evidence	28
1.3.3 Contradicting Evidence	33
1.3.4 Summary	38
1.3.5 Research Question	40
1.4 Methodology: Possibilities of Studying the Mapping from Sound to Representation	43
1.4.1 Electroencephalography - EEG	43
1.4.1.1 Neural & Electric Basis of the EEG	43
1.4.1.2 Measuring EEG & Event-Related Brain Potentials	44
1.4.2 ERP Components	47
1.4.2.1 Automatic Change Detection: The MMN	47
1.4.2.1.1 MMN Generation	50
1.4.2.1.2 MMN & Memory	51
1.4.2.1.3 MMN & Language Processes	54
1.4.2.1.4 Variations of the Oddball Paradigm	57
1.4.2.1.5 The MMN as Difference Response Controlling the Effects of Physical Stimulus Differences	59
1.4.2.2 Stimulus Updating: The P300	61

1.4.2.3	MMN & P300	64
2	General Methods	67
2.1	Participants	67
2.2	Stimuli	67
2.3	Experimental Procedure	68
2.4	Data Acquisition and Analysis	69
3	Experiments on Place- and Manner-of-Articulation Features	71
3.1	Processing & Representation of Place-of-Articulation Features	71
3.1.1	Experiment 1: Underspecification of the Place Feature [CORONAL] in Vowels	71
3.1.1.1	Methods	74
3.1.1.2	Results	79
3.1.1.3	Discussion of Experiment 1	82
3.1.2	Experiment 2: Underspecification of the Place Feature [CORONAL] in Stop Consonants	87
3.1.2.1	Methods	89
3.1.2.2	Results	93
3.1.2.3	Discussion of Experiment 2	96
3.2	Processing and Representation of Manner-of-Articulation Features	100
3.2.1	Experiment 3: Mental Representations of Place- & Manner-of-Articulation Features	100
3.2.1.1	Methods	102
3.2.1.2	Results	106
3.2.1.3	Discussion of Experiment 3	109
3.2.1.4	General Discussion of Place Features for Stops & Manner Features	113
3.2.2	Experiment 4: Underspecification of the Manner Feature [PLOSIVE]	117
3.2.2.1	Methods	120
3.2.2.2	Results	123
3.2.2.3	Discussion of Experiment 4	126

3.2.3	Experiment 5: Mental Representations of Place- and Manner- of-Articulation Features in Different Speech Sounds	132
3.2.3.1	Methods	135
3.2.3.2	Results	139
3.2.3.3	Discussion of Experiment 5	142
3.2.4	Experiment 6: Reaction Time Experiment	150
3.2.4.1	Methods	153
3.2.4.2	Results	155
3.2.4.3	Discussion of Experiment 6	156
4	General Discussion	159
4.1	Possible Influencing Aspects	168
4.1.1	Latency Effects	168
4.1.2	Acoustics	170
4.1.3	Frequency Effects	171
4.1.4	P300 Effects – Automatic Attention Switch	172
4.2	Comparison with Predictions of Alternative Models	173
4.3	Questions, Solutions, and Ideas	180
4.3.1	Open Questions	180
4.3.2	What is Novel about this Research?	182
4.3.3	Future Outlook, Further Research Questions	182
4.4	Conclusion	184
5	Appendices	186
5.0	Figure & Table Index	186
5.1	Figures	189
5.1.1	Experiment 1: ‘Vowels’: Pitch & Intensity Contours	189
5.1.2	Experiment 2: ‘Plosives’: Pitch & Intensity Contours	191
5.1.3	Experiment 1: Standards, Deviants, & MMN Waveforms	192
5.1.4	Experiment 2: Standards, Deviants, & MMN Waveforms	196
5.1.5	Experiment 3: Standards, Deviants, & MMN Waveforms	199
5.1.6	Experiment 4: Standards, Deviants, & MMN Waveforms	201
5.1.7	Experiment 5: Standards, Deviants, & MMN Waveforms	204
5.2	Statistics	209
5.2.1	Experiment 1: Standard/Deviant Analysis	209

5.2.2	Statistical Tables	212
5.2.2.1	Experiment 1: Place Features ‘Vowels’ (Tables 14a-d)	212
5.2.2.2	Experiment 2: Place Features ‘Plosives’ (Tables 15a-d)	216
5.2.2.3	Experiment 3: Place & Manner Features (Tables 16a-c)	220
5.2.2.4	Experiment 4: Manner Features (Tables 17a-c)	222
5.2.2.5	Experiment 5: Place & Manner Features (Tables 18a-c)	224
5.2.2.6	Experiment 6: Reaction Times (Table 19)	228
6	References	229

Glossary (Abbreviation Index)

ANOVA	analysis of variance
BESA	brain electric source analysis software
CV	consonant-vowel
CVC	consonant-vowel-consonant
CVC̃	consonant-nasalized vowel-consonant
CVCCVC	consonant-vowel-consonant-consonant-vowel-consonant
CVN	consonant-vowel-nasal
CNV	contingent negative variation
Cz	central electrode
dB _{SPL}	decibel <small>sound pressure level</small>
DC	direct current
DOC	direction of change
DP	distraction potential
EEG	electroencephalogram or electroencephalography
(E)LAN	(early) left anterior negativity
EOG	electro-oculogram
EROS	event-related optical signals
ERP	event-related brain potentials
F0	fundamental frequency
F1	first formant frequency
F2	second formant frequency
F3	third formant frequency
fMRI	functional magnetic resonance imaging
FUL	Featurally Underspecified Lexicon Model (Lahiri & Reetz, 2002)
Fz	fronto-central electrode
Hz	Hertz
iMMN	identity MMN
ITI	inter stimulus interval
kHz	kilohertz
kΩ	kiloohm
LPC	late positive component
μV	microvolts
MEG	magnetencephalogram or magnetencephalography
MMF	magnetic mismatch field
MMN	mismatch negativity
ms	milliseconds
n	number of subjects/trials/samples
n.s.	not significant
PET	positron emission tomography
POI	pair of inversion
Pz	parieto-central electrode
RMS	root-mean-square power
RON	reorienting negativity
RT	reaction time
SEM	standard error of the mean
STG	superior temporal gyrus
TRACE	TRACE model of speech perception (McClelland & Elman, 1986)
VC	vowel-consonant
VOT	voice-onset-time
/x/	sounds between slashes as in /x/ mark the underlying representation of a sound stored in the mental lexicon, i.e. activated by the standard.
[x]	sounds between square brackets as [x] represent the phonetic surface form of the acoustic incoming sound, i.e. triggered by the deviant.

Figure Legend

- Fig. 1: A distributed model of speech perception (Gaskell & Marslen-Wilson, 1997, 1999)
- Fig. 2: Feature organization of the FUL model (Lahiri & Reetz, 2010)
- Fig. 3: MMN waveforms for all vowel pairs reversed as standard & deviant (Eulitz & Lahiri, 2004)
- Fig. 4: ERPs to standard tones & deviant tones of different frequencies (Sams et al., 1998a)
- Fig. 5: Diagram of the two oddball paradigms, the classic & roving-standard design
- Fig. 6: Experiment 1: 'Vowels' Oscillograms of acoustic stimuli of words & nonwords
- Fig. 7: Experiment 1: 'Vowels': Acoustic characteristics of the three vowel categories
- Fig. 8: Experiment 1: 'Vowels': Grand-average waveforms for the non-conflicting & the conflicting pair of inversion for words & nonwords
- Fig. 9: Experiment 1: 'Vowels': Bar graphs of the MMN peak amplitudes grand for all conflicting & non-conflicting conditions
- Fig. 10: Experiment 2: 'Plosives': Oscillograms of acoustic stimuli of CV-syllables with plosives
- Fig. 11: Experiment 2: 'Plosives': Grand-average difference waveforms of all six conditions the [Labial]~[Coronal], [Labial]~[Dorsal], & [Coronal]~[Dorsal] pair of inversion
- Fig. 12: Experiment 2: 'Plosives': Bar graphs of the MMN mean amplitudes of all six conditions for the early & late MMN time window
- Fig. 13: Experiment 3: Place & Manner Features: Oscillograms of acoustic stimuli of the four VCV-syllables
- Fig. 14: Experiment 3: Place & Manner Features: Pitch and intensity contours of the VCV-syllables
- Fig. 15: Experiment 3: Place & Manner Features: Grand-average difference waveforms of all four conditions, the PLACE CHANGE, & the MANNER CHANGE pair of inversion
- Fig. 16: Experiment 3: Place & Manner Features: Bar graphs of the MMN mean peak amplitudes of all four conditions
- Fig. 17: Experiment 4: Manner Features: Pitch and intensity contours of the VCV-syllable contrasts
- Fig. 18: Experiment 4: Manner Features: Grand-average waveforms for all six conditions, the Manner Change Asymmetric, the Manner Change Symmetric, and the Manner & Place Change Symmetric pair of inversion
- Fig. 19: Experiment 4: Manner Features: Bar graphs of the MMN mean peak amplitudes of all six conditions, the Pair of Inversion Manner Change Asymmetry, the control condition Manner Change Symmetry, & the Pair of Inversion Manner & Place Change Symmetry
- Fig. 20: Experiment 5: Place & Manner Features: Oscillograms of acoustic stimuli of the four VCV-syllables
- Fig. 21: Experiment 5: Place & Manner Features: Pitch and intensity contours of the VCV-syllable contrasts
- Fig. 22: Experiment 5: Place & Manner Features: Grand-average waveforms for all eight conditions, the PLACE contrasts with nasals & plosives, & the MANNER contrasts with coronals & labials

Fig. 23: Experiment 5: Place & Manner Features: Bar graphs of the MMN mean amplitudes & MMN peak latencies of all eight conditions

Fig. 24: Experiment 6: Reaction-Times: Bar graphs of the mean reaction times for all sound contrasts, the place & manner-of-articulation features

Fig. 25: Experiment 6: Reaction-Times: Bar graphs of the mean amplitudes of the P3a components for all conditions across Experiments 3-5

Table Legend

Table 1: The ternary matching logic of the FUL model (Lahiri, 1999)

Table 2: Experiment 1: ‘Vowels’: Predictions about the amount of feature conflict for the vowel contrasts of models assuming full specification & underspecification

Table 3: Experiment 1: ‘Vowels’: Log-values of frequency counts of phonemes & phoneme sequences

Table 4: Experiment 1: ‘Vowels’: German words and nonwords used in the eight experimental conditions

Table 5: Experiment 2: ‘Plosives’: Predictions about the amount of feature conflict for the CV-contrasts used in the six experimental conditions

Table 6: Experiment 2: ‘Plosives’: Log-values of frequency counts of phonemes & phoneme sequences of the CV-syllables

Table 7: Experiment 3: Place & Manner Features: Predictions about the amount of feature conflict for the nonword stimuli used in the four experimental conditions

Table 8: Experiment 3: Place & Manner Features: Log-values of frequency counts of phonemes & phoneme sequences of the VCV-syllables

Table 9: Experiment 4: Manner Features: Predictions about the amount of feature conflict for the nonword stimuli used in the six experimental conditions

Table 10: Experiment 5: Place & Manner Features: Predictions about the amount of feature conflict for the nonword stimuli used in the eight experimental conditions

Table 11: Experiment 5: Place & Manner Features: Log-values of frequency counts of phonemes & phoneme sequences of the VCV-syllables

Table 12: Experiment 6: Reaction-Times: Predictions for the reaction times following MMN latencies or P3a amplitudes

Table 13: Overview of all experimental contrasts fulfilling the predictions of the different models

0 Abstract

To date it is unclear how the brain deals with the enormous variation in speech and how words are recognized with such varying input. It also remains unknown what exactly is stored in the mental lexicon and how detailed phonological information must be represented to ensure fast and adequate speech comprehension and processing. Different kinds of psycholinguistic models have been proposed in dealing with such variation and how speech sounds are stored in the mental lexicon. This dissertation aims to resolve this question by using electrophysiological methods in the investigation of the neuronal representation of phonological information.

Chapter 1 gives an introduction to four different models of speech perception and representation. The account, which is tested in this research – the Featurally Underspecified Lexicon Model (FUL; Lahiri & Reetz, 2002, 2010) – assumes abstract and underspecified representations of phonological features. When dealing with variation, all models except for the FUL model assume full specification of lexical form that relies on experience or context. The Mismatch Negativity, an automatic change detection response in the brain, has been shown to be a powerful tool in this line of research as it enables the investigation of the representation of specific phonological features by contrasting speech sounds. Five MMN experiments were set up to test predictions of the FUL model in contrast to other models that differ in the degree to which phonological feature information (i.e. place and manner of articulation) is specified or underspecified. Only models assuming underspecified phonological features such as FUL predict an asymmetry of MMNs within the reversal of sound contrasts, presented as standard and deviant. For instance, sounds that are underspecified regarding a certain feature, cannot build up a representation when presented as standard. Consequently they do not conflict with the incoming deviant sound, which is reflected in reduced MMN amplitudes. In the reversed case, when a sound is assumed to be specified for a specific feature, it will be pre-activated by the standard. Therefore a conflict occurs with other incoming deviant sounds, which will elicit larger MMN responses. Models proposing full storage of all phonological information would predict equal MMNs within the reversal of sound contrasts.

Chapter 2 presents the general methods used in all of the altogether six experiments. Adult subjects were presented with speech sounds embedded into words, nonwords, and syllables that were contrasted, each serving as standard and deviant in an MMN design.

In Chapter 3 the studies with their results are presented and discussed. More specifically, Experiment 1 tested the FUL model's assumptions of the underspecified place feature [CORONAL] in vowels. Three German vowels were placed in words and identical nonwords each serving as standard and deviant and thereby comparing the place features

[CORONAL] and [DORSAL]. Results showed MMN asymmetries with the largest amplitudes when the unspecified coronal deviant mapped onto the represented dorsal standard, which induced a conflict. This effect was strongest for nonwords where additional lexico-semantic ambiguities did not influence the phonological effects.

Experiment 2 attempted to expand the findings for underspecified place-of-articulation features to other speech sounds, using three syllable initial stop consonants as standards and deviants, and contrasting the features [LABIAL], [CORONAL], and [DORSAL]. Unexpectedly, the MMN results were rather weak. The plosive sound contrasts did not elicit reliable MMNs for either contrast possibly not having tapped into lexical representations. Considerations about these non-results are discussed.

The following three MMN experiments (3-5) were aimed at generalizing the underspecification account to other feature dimensions such as manner of articulation. Consonants embedded in VCV-syllables were differently contrasted. Experiment 3 made the first step into another feature dimension by teasing apart MMN effects by systematically manipulating two featural dimensions, place ([CORONAL] ~ [DORSAL]), and manner of articulation ([NASAL] ~ [STRIDENT]). The pattern of results was compatible with the idea that abstract phonological feature information activates specific memory traces and does not activate underspecified information. Furthermore, the phonetic contrasts evoked MMNs, which differed in the two feature dimensions under investigation. It was concluded that differences in manner are perceptually more salient than differences in place of articulation.

Experiment 4 focused on manner-of-articulation features. Asymmetric MMNs were predicted for the reversal of the [PLOSIVE] ~ [NASAL] contrast due to the assumed underspecification of the feature [PLOSIVE], whereas the equally specified [NASAL] ~ [STRIDENT] reversals were predicted to evoke symmetric MMNs. The observed pattern of MMN differences supported predictions assuming underspecified phonological representations for the feature [PLOSIVE]. A third condition explored different feature types within one contrast, comparing manner features in one direction of standard-deviant change, and place features in the reversed condition. The observed MMNs showed enhancement for the manner conflicting case. These results seem to support the idea of a hierarchical structure within the featural dimensions during speech perception, with manner of articulation being superior over place in the extraction of feature information.

Experiment 5 extended the assumptions made for underspecified manner and place features by investigating additional speech sounds such as nasals, plosives, coronals, and labials. The expected asymmetric MMN activations were found for the conflicting ~ non-conflicting place contrasts for nasals and plosives. For the manner contrasts asymmetric patterns were only found for the coronal sounds.

Experiment 6 was a behavioral discrimination study to control for acoustic and attentional effects of all the consonantal sound contrasts used in Experiments 3-5. The results showed similar reaction times between the acoustically very close place contrasting sounds such as plosives and nasals. Reaction times differed between the manner contrasts [n] ~ [d] and [n] ~ [z]. These results can be related to attentional effects of the acoustically more salient sounds. Additionally, the reaction time results seemed independent of the MMN results, in that they reflected only acoustic differences, whereas the MMN also reflects representational differences.

Breaking new ground by looking at different feature dimensions in this dissertation such as place and manner of articulation with different speech sounds, provides further support for the notion that mental representations of phonological features are not a one-to-one relation between the acoustic speech signal and their mental representations. Rather phonological information may be more abstract compared to theories suggesting full storage of all phonological details. These reported MMN asymmetries between conditions for the same acoustic contrasts in Chapter 3 clearly suggest that the brain refers to underspecified phonological representations as a basic principle in the functional organization of the mental lexicon during speech perception.

Chapter 4 gives a general discussion of all results, discussing them in the light of different psycholinguistic models and further influencing factors such as acoustics and frequency effects.

0.1 Zusammenfassung

Wie das Gehirn mit der großen Varianz in der Sprache umgeht und wie Wörter aus dem variantenreichen Sprachsignal erkannt und verstanden werden, ist bisher ungeklärt. Es ist auch unklar, was genau im mentalen Lexikon abgespeichert ist und wie detailliert phonologische Informationen repräsentiert sein muss, um eine schnelle und angemessene Sprachverarbeitung und Sprachperzeption zu gewährleisten. Verschiedene psycholinguistische Modelle wurden entwickelt, die mit diesen sprachlichen Variationen umgehen und die zu erklären versuchen wie sprachliche Laute im mentalen Lexikon abgespeichert sein könnten. Diese Dissertation hat zum Ziel, der Frage nach der neuronalen Repräsentation phonologischer Information mit Hilfe von elektrophysiologischen Methoden nachzugehen.

Kapitel 1 stellt vier verschiedene Modelle der Sprachperzeption und -repräsentation vor. Das Modell das in dieser wissenschaftlichen Arbeit getestet wird – das Featurally Underspecified Lexicon Model (FUL; Lahiri & Reetz, 2002, 2010) – nimmt abstrakte und unterspezifizierte Repräsentationen von phonologischen Merkmalen an. Beim Umgang mit Sprachvariation nehmen alle Modelle außer FUL eine volle Spezifikation lexikalischer Formen an, die auf Erfahrung oder kontextueller Information basieren. Die Mismatch Negativity, eine automatische Hirnreaktion, die akustische Veränderungen detektiert, hat sich für diese Art von wissenschaftlicher Forschung als effektive Methode gezeigt. Sie ermöglicht die Erforschung von Repräsentationen bestimmter phonologischer Merkmale, indem Sprachlaute kontrastiert werden. Fünf MMN-Experimente wurden durchgeführt, um Annahmen von FUL im Vergleich zu anderen Lexikon Modellen zu testen, die sich darin unterscheiden, welche Art von phonologischer Merkmalsinformation (z.B. Artikulationsort und -art) spezifiziert oder unterspezifiziert ist. Nur Modelle, die unterspezifizierte phonologische Merkmale annehmen, erwarten MMN-Asymmetrien in der Umkehr von Lautkontrasten. So können zum Beispiel Laute, die hinsichtlich eines bestimmten Merkmals unterspezifiziert sind keine Repräsentation aufbauen, wenn sie als Standard präsentiert werden. Sie konfliktieren dann nicht mit dem eingehenden devianten Laut, was sich in einer geringen MMN widerspiegeln sollte. In der umgekehrten Richtung, d.h. wenn ein Laut hinsichtlich eines bestimmten phonologischen Merkmals spezifiziert ist, sollte dieses Merkmal vom Standard voraktiviert werden. Damit entsteht ein Konflikt mit anderen eingehenden abweichenden Lauten. Dies sollte sich in einer erhöhten MMN widerspiegeln. Phonologische Modelle, die die volle Spezifikation phonologischer Repräsentationen annehmen, erwarten gleiche MMNs zwischen kontrastierenden Lauten, unabhängig davon welcher als Standard oder Deviant präsentiert wird.

Kapitel 2 stellt die allgemeinen Methoden vor, die in allen insgesamt sechs Experimenten durchgeführt wurden. Erwachsenen Versuchspersonen wurden Sprachlaute präsentiert, die in Wörter, Pseudowörter und Silben eingebettet waren. Sie wurden kontrastiert, d.h. sie dienten jeweils als Standard und als Deviant in einem MMN-Design.

In Kapitel 3 werden Studien und ihre Ergebnisse präsentiert und diskutiert. Insbesondere Experiment 1 testete die Annahmen des FUL-Modells bezüglich der Unterspezifikation des Artikulationsortsmerkmals [CORONAL] in Vokalen. Drei deutsche Vokale wurden in Wörter und äquivalente Pseudowörter eingebettet, die jeweils als Standard und Deviant dienten und dabei die Merkmale des Artikulationsortes [CORONAL] und [DORSAL] kontrastierten. Die Ergebnisse zeigten asymmetrische MMNs, mit einer erhöhten Amplitude für die konfligierende Bedingung, wenn der unterspezifizierte coronale Deviant auf einen repräsentierten dorsalen Standard traf. Dieser Effekt war für die Pseudowörter am stärksten, bei denen zusätzliche lexikalische Ambiguitäten die phonologischen Effekte nicht weiter beeinflussten.

Experiment 2 versuchte die Ergebnisse der unterspezifizierten Merkmale des Artikulationsortes auf andere Sprachlaute zu erweitern, indem drei silbeninitiale Stop-Konsonanten als Standard und Deviant verwendet und damit die Merkmale [LABIAL], [CORONAL] und [DORSAL] verglichen wurden. Unerwarteterweise, zeigten sich die MMN-Ergebnisse als ziemlich schwach. Die Plosive lösten für keinen Lautkontrast reliable MMNs aus, was möglicherweise daran lag, dass die lexikalischen Repräsentationen nicht richtig aktiviert wurden. Verschiedene Überlegungen bezüglich dieser Ergebnisse werden in Betracht gezogen und diskutiert.

Die folgenden drei Experimente (3-5) zielten darauf ab, die Annahmen der Unterspezifikation auf andere Merkmalsdimensionen wie auf Merkmale der Artikulationsart zu generalisieren. Konsonanten waren eingebettet in VCV-Silben und wurden unterschiedlich kontrastiert. Experiment 3 machte einen ersten Schritt in eine andere Merkmalsdimension, indem MMN-Effekte zerlegt und systematisch zwei Merkmalsdimensionen, wie Artikulationsort ([CORONAL] ~ [DORSAL]) and Artikulationsart ([NASAL] ~ [STRIDENT]) variiert wurden. Die Ergebnisse stimmen mit der Idee überein, dass abstrakte phonologische Merkmalsinformationen spezifische Gedächtnisspuren aktivieren und unterspezifizierte Information nicht aktiviert werden. Zusätzlich lösten die phonetischen Kontraste MMNs aus, die sich in den beiden untersuchten Merkmalsdimensionen unterschieden. Es wurde vermutet, dass die Unterschiede in der Artikulationsart perzeptuell salienter sind als die Unterschiede im Artikulationsort.

Experiment 4 fokussierte sich auf Merkmale der Artikulationsart. Aufgrund der angenommenen Unterspezifikation des Merkmals [PLOSIV], wurden MMN-Asymmetrien für den [PLOSIV] ~ [NASAL] Kontrast erwartet, hingegen für die gleichermaßen spezifizierten

[NASAL] ~ [STRIDENT] Merkmalskontraste wurden symmetrische MMNs angenommen. Die beobachteten MMN-Unterschiede sprechen für Annahmen der Unterspezifikation von phonologischen Repräsentationen des Merkmals [PLOSIV]. Eine dritte Bedingung untersuchte unterschiedliche Merkmalskonflikte innerhalb eines Lautkontrastes. Es wurden Merkmale der Artikulationsart in einer Richtung der Standard-Deviant-Präsentation mit Merkmalen des Artikulationsortes in umgekehrter Standard-Deviant-Richtung verglichen. Die aufgezeichneten MMNs zeigten eine erhöhte Amplitude für die Merkmalskonflikte der Artikulationsart. Diese Ergebnisse stützen die Idee einer hierarchischen Struktur innerhalb der verschiedenen Merkmalsdimensionen bei der Sprachperzeption, bei der die Merkmale der Artikulationsart bei der Extraktion von Merkmalsinformationen über denen der Merkmale des Artikulationsorts stehen.

Experiment 5 erweiterte die Annahmen, die für unterspezifizierte Artikulationsart und -ortsmerkmale gemacht werden, indem weitere Sprachlaute, wie nasale, plosive, coronale und labiale Laute, untersucht wurden. Die erwarteten MMN-Asymmetrien wurden für die konfligierenden ~ nicht-konfligierenden Artikulationsortskontraste für Nasale und Plosive gefunden. Für die Kontraste der Artikulationsart wurden diese Asymmetrien nur für die coronalen Laute gefunden.

Experiment 6 war eine behaviorale Diskriminationsstudie, die akustische und aufmerksamkeitsbezogene Prozesse aller konsonantischen Lautkontraste, die in den Experimenten 3-5 untersucht wurden, kontrollieren sollte. Die Ergebnisse zeigten ähnliche Reaktionszeiten für die akustisch sehr ähnlichen Laute, die den Artikulationsort kontrastierten wie Nasale und Plosive. Reaktionszeiten unterschieden sich zwischen den Kontrasten der Artikulationsart [n] ~ [d] and [n] ~ [z]. Diese Ergebnisse können auf Aufmerksamkeitseffekte hinsichtlich der akustisch salienteren Laute bezogen werden. Weiterhin schienen die Reaktionszeiten unabhängig von den MMN-Ergebnissen zu sein, dahingehend, dass sie nur akustische Unterschiede widerspiegeln, die MMN hingegen aber auch noch repräsentationale Unterschiede reflektiert.

Indem erstmalig verschiedene Merkmalsdimensionen in dieser Dissertation untersucht wurden, wie Merkmale des Artikulationsortes und der Artikulationsart mit unterschiedlichen Sprachlauten, ist weitere Evidenz für die Sichtweise geschaffen worden, dass mentale Repräsentationen phonologischer Merkmale keine eins-zu-eins Übersetzung zwischen dem akustischen Sprachsignal und der mentalen Repräsentation sind. Vielmehr scheint die phonologische Information mehr abstrakt repräsentiert zu sein als Theorien, die eine volle Speicherung aller phonologischen Details annehmen, postulieren. Die berichteten MMN-Asymmetrien in Kapitel 3 zwischen Bedingungen mit gleichen akustischen Kontrasten weist darauf hin, dass das Gehirn unterspezifizierte phonologische

Repräsentationen als grundlegendes Prinzip in der funktionalen Organisation des mentalen Lexikons während der Sprachwahrnehmung nutzt.

Kapitel 4 diskutiert alle Ergebnisse im Licht der verschiedenen psycholinguistischen Modelle unter Berücksichtigung weiterer beeinflussender Faktoren wie Effekte der Frequenz oder der akustischen Eigenschaften der Laute.

1 General Introduction

As language listeners, we easily process words, seemingly without any effort. Spoken language comprehension is among the most complex human abilities. In fact, a complex cascade of processes underlies our human capacity to understand spoken language. Numerous brain areas work together to analyze the acoustic information that is transient and often ambiguous, selecting the proper words by mapping the auditory sensory input onto stored lexical knowledge, extracting the meaning of these words and integrating them into a continuing sentential or discourse context, all of which happens within only a few hundred milliseconds (Marslen-Wilson & Tyler, 1980; Norris, 1994; cited in Hagoort, 2008).

1.1 Process of Word Recognition

A central part of speech perception is word recognition. If a listener hears [pʊsɪ ɡəlʊə pʊld ðə hæmbre:k əv ðə æstən ma:tin di: bi: faɪv æz ha:d æz ʃɪ kʊd waɪl ðə hænsʌm dʌbl əʊ e:dʒnt græbd hɪz hæŋɡʌn ɪn əʊdə] ‘*Pussy Galore pulled the handbrake of the Aston Martin DB5 as hard as she could while the handsome double-0 agent grabbed his handgun in order....*’ the only way for the listener to understand this sentence is to map the acoustic-phonetic information of the speech signal onto stored knowledge about the sound forms of words. This process is complicated by the variable nature of spoken language and still remains unresolved.

Most models of word recognition agree that the comprehension of spoken words involve three major components, or processing levels. These include an early auditory/acoustic analysis, followed by a lexical, and a semantic processing level (Ellis & Young, 1988; Lesser & Milroy, 1993; cited in Pettigrew, Murdoch, Chenery & Kei, 2004). According to these models, the first stage of auditory word comprehension is performed by an analysis in the auditory system, which identifies certain units (e.g. phonemes, phonological features, etc.) within the incoming acoustic waveform. The results of this analysis are transferred to the auditory input lexicon, which contains the ‘store’ of known words and their characteristics. Ultimately, a match is sought between the speech input and stored entries in the lexicon. If a match is found, the appropriate lexical entry is activated, which in turn activates the representation of the spoken word’s meaning, located in the semantic system. Given that spoken word recognition is essential for bridging the gap between sound and meaning, it is important to consider how models of spoken word comprehension accommodate the process of recognition. Frauenfelder and Tyler (1987), for instance, identified several stages of lexical processing, including a *contact*, *activation*,

selection, and *access* stage. During the initial *contact* stage, the listener takes the speech wave as sensory input and generates representations that contact the form-based representations stored internally which are associated with each lexical entry. The next stage engages the *activation* of the lexical items that refer to those lexical entries matching the contact representation to some degree during the initial contact phase, and may then change in state. This activation stage is then followed by the *selection* of the appropriate lexical item from a set of activated candidates. Following the initial contact and activation of a subset in the lexicon, accumulating sensory input continues to map onto this subset of entries until the intended lexical entry is *selected* and recognized. It is suggested, that word recognition occurs at the end of the selection stage, when the listener has determined which lexical item was heard. At the end of this process, the whole information listed in the lexical entry is *accessed*, such as semantic and syntactic information, and the listener identifies and recognizes words. *Lexical access* therefore refers to the point at which various properties of stored lexical representations (phonological, syntactic, semantic, etc.) become available to the listener. In other words, lexical access refers to the initial mapping of the incoming acoustic signal onto lexical representations. To fully understand a heard message, further integration processes need to take place, where the retrieved information is integrated into higher-level context representations of the preceding part of an utterance. Word recognition occurs when a single lexical entry is selected.

Although this framework attempts to clarify the stages of word recognition in auditory processing, it remains unclear how exactly the extraction of phonetic information from the speech signal (e.g. as phonetic features, acoustic properties) is transmitted into representations to access words and meanings.

The physical speech signal that makes up words and sentences consists of a stream of acoustic energy, varying over time in amplitude and spectral shape, involving different frequencies. Speech sounds and spoken words are not invariant and stable. There is a huge variability within and across speakers. Many relevant fine distinctions can be ambiguous by the kinds of phonological processes that commonly take place during speech production, such as assimilation, deletion and insertion of sounds, shortening and lengthening.

Lexical activation seems fine-tuned to the acoustic-phonetic detail in the speech signal. Variation in the speech signal can influence lexical activation to different degrees, but in most cases the listener seems to easily deal with it. The phonological context, for instance, can crucially determine the activation of words. The word *hand* can be pronounced in different ways, depending on the following context. In compounds like *handbrake*, or

handgun or words such as *handsome*, the original *hand* will result in [haŋ] as in **hangun*¹ [ham] in **hambag* or [han^d] in *handsome*. This assimilated information appears to influence lexical activation. While context assimilation describes a process across words, a similar viable within-word process is co-articulation (segmental overlap) – the vocal tract producing a sound is already preparing for the following sound thereby influencing the preceding and following sounds. Furthermore, the speaker’s gender, age, dialectal variations, speaking rate and style, as well as background noise, and the position and density of words in the prosodic structure (McQueen, 2004) all lead to variations in the speech signal.

The process of assimilation is a very common phenomenon that occurs during the unfolding of the speech signal. Assimilation means the alignment of neighboring segments in at least one phonetic parameter. It is a language-specific process that causes a discrete change from one phoneme to another. Assimilations are determined by the direction in which they operate, by the phonetic parameter that is affected, and the temporal dimension. During *progressive assimilation* the following phoneme is influenced (e.g. *haben* ‘to have’ [bn] → [bm]), whereas in *regressive assimilation* the previous sound is influenced (e.g. *Eisschrank* ‘refrigerator’ [sʃ] → [ʃʃ]). Assimilations in both directions are possible as well (e.g. *haben* ‘to have’ [bn] → [bm] → [m]). The change can refer to the place of articulation (e.g. *Wein Bar* ‘wine bar’ [nb] → [mb]), to manner of articulation (e.g. *Signal* ‘signal’ [gn] → [ŋn]), or voicing. Assimilation can also reduce forms in the case of voiced plosives preceding a nasal, where only nasals remain (e.g. *finden* ‘to find’ [nd] → [nn]; *Agnes* [gn] → [ŋn]; *Bombe* ‘bomb’ [mb] → [mm]). They can occur in different positions depending on the linguistic unit such as at word –, morpheme –, or syllable boundaries (e.g. in compounds as in *handgun* or *handbag* [dg] → [ŋg] or [db] → [mb]) as well as within syllables, also depending on the context (Kohler, 1995).

Assimilation of place of articulation often affects word final coronal consonants such as /d/, /t/, and /n/, which in connected speech can take on the place of articulation of the neighboring labial or velar consonant, as in the example of *handbag* resulting in **hambag*. These changes may result in completely assimilated sounds with no residual information of the original segment (Zimmerer, Reetz, Lahiri, 2009). In other cases, they may be only partial, where a segment is produced with phonetic information relating to two different places of articulation (Nolan, 1992; Gow, 2002, 2003). Importantly, these place assimilation processes are *asymmetric* in that coronal consonants can change to labial or velar place of articulation in the following segment, whereas the reverse case is hardly ever observed (Paradis & Prunet, 1991; Lahiri & Reetz, 2002). Labials and velars hardly ever assimilate to another place. For instance words ending with a labial /m/ as in *warm* will not change to a

¹ The ‘asterisk’ symbol [*] refers to a nonword.

dorsal [ŋ] in **war[ŋ] glove* or to a coronal [n] in **war[n] day*. Only coronal sounds are special in that they frequently undergo assimilation. They are assumed to have a special status. Paradis and Prunet (1991) explain this asymmetry of place assimilation in terms of a default status of coronals, which means they do not contain specific lexical representations. They are underspecified for place of articulation. Similarly, assimilation of manner of articulation often affects plosive consonants, in particular when they co-occur with nasal consonants. [PLOSIVE + NASAL] sequences are rather rare in languages, and often become entirely nasal like in *bomb*, *lamb* ([mb] → [m]), *painting* ([nt] → [n]) and *kindergarden* ([nd] → [n]), but do not assimilate to the reversed. It could be that the feature [PLOSIVE] also is underspecified as it is very likely to undergo change, similar to the feature [CORONAL]. However this is still unknown territory.

These asymmetries lead to variances in the acoustic speech signal that are handled by listeners without any difficulty. But how do these variants (often resulting in nonwords) still activate the correct words in the lexicon? Different models of speech perception propose different mechanisms and views on how the perceptual system deals with this variation in order to map the acoustic signal onto the correct lexical representation. For instance, some propose detailed memories about variants stored in memory, others assume statistical computations and learnt probabilities about variant speech sounds, or infer context-sensitive inference mechanisms to deal with variation and this mapping process. Yet others assume abstract and sparse lexical representations are stored in the lexicon, where not all information is stored. This last model refers to the theory of underspecification which assumes that some feature specifications are absent in the underlying lexical representation (Archangeli, 1988; cited in Paradis & Prunet, 1991; Lahiri, 1999; underspecification will be discussed in more detail in the following chapter and work), which makes the spreading of specified features such as noncoronals (labials and velars) onto an unspecified coronal place possible.

A listener is faced with the challenge of mapping these variances of the spoken speech stream for example caused by assimilation onto their stored representations and knowledge of spoken words. The number of different acoustic realizations in which a particular speech sound or a word is manifested is potentially infinite. Therefore, a one-to-one mapping of acoustic realization to the sound identity in the lexicon seems problematic, although proposed by some models (e.g. Bybee, 2001).

There is still disagreement about the nature of the processes and representations² that this mapping may involve. It is necessary to perform an acoustic-phonetic analysis of the speech signal, translating this varying signal into representations of the sounds prior to lexical access in an initial *contact* stage. The activation of a word is modulated by several factors, such as its own goodness of fit to its corresponding lexical representation, but also by the number of other words that have a potentially better fit. In particular, the quality of match between the candidate word and the speech input is again determined in terms of segmental information (such as phonemic or featural information), as well as in terms of suprasegmental information (e.g. stress pattern). Finally, the match depends on the fitting of the word's form in the phonological context (McQueen, 2004). The better the fit, the greater the activation of the target will be. Vice versa, the poorer the fit, the less activation a lexical item will receive. In other words, the success of the lexical access depends crucially on the goodness-of-fit between the sensory information and stored information about the form of words.

Psycholinguistic models differ with respect to these different word recognition processes, for instance in terms of the number and kinds of levels of representation (varying from pre-lexical to lexical), and the kind of information that is extracted from the acoustic signal represented in the mental lexicon (e.g. spectral sequences, Klatt, 1989; phonetic features, Stevens, 1986, 2002, Stevens & Keyser, 2008; phoneme units, Pisoni & Luce, 1987; according to Frauenfelder & Tyler, 1987, these *contact representations* all take place in the *initial lexical contact phase*). Also, the mechanisms of competition between the activated items (graded or probabilistic vs. categorical) differ between the models, as well as the consequences of a mismatch between the input and the representation influencing the selection process (where re-analysis of the sensory input may be required). Furthermore, influences of semantic, syntactic, contextual information during lexical access, and flow of information as bottom-up or top-down connections (autonomous vs. interactive views; feed-forward vs. feed-back models), are taken into account differently (for review see McQueen, 2004).

² There is evidence that the primary substrate for constructing sound-based representations of speech involves auditory cortices in the posterior superior temporal lobe bilaterally (spectro-temporal analysis and phonological processing). From this point on in the recognition process it is suggested that there is a divergence in the processing stream, with one stream involving more ventral-posterior structures (ventral pathway; anatomical structures in the superior and middle portions of the temporal lobe), and the other stream involving more dorsal-anterior systems (dorsal pathway; inferior parietal and posterior frontal systems), both being more lateralized to the left hemisphere. The ventral pathway seems to be important for interfacing sound-based representations of speech with distributed conceptual representations linked to the mental lexicon (speech recognition). The dorsal pathway plays a greater role in accessing certain sub-lexical speech segments translating acoustic speech signals into articulatory representations in the frontal lobe. A variety of evidence suggests that the left and right systems make different contributions to the perceptual analysis of speech in a multi-time resolution process integrating information on different timescales (Hickock & Poeppel, 2000, 2007).

Moreover, many models of speech perception dealing with this encoding issue differ with respect to how the speech signal is mapped onto the mental representation in order to access words. There is disagreement in the literature about the nature of this process and even about the nature of the representations. Speech perception must transform the acoustic/phonetic input signal, as a continuously varying waveform into some kind of discrete mental representation for successful comprehension interfacing auditory information with lexical representation. These perception models vary from highly abstract lexical representations, whereas only distinctive, non-redundant information is stored and the speech signal is converted into a set of context-independent abstract phonological features (e.g. Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002; Pallier, Colomé, & Sebastián-Gallés, 2001), to completely stored features, phonemes or phoneme-strings (McClelland & Elman, 1986; Norris, 1994). Others store extremely detailed, episodic acoustic representations where all variant pronunciations and a complete collection of memory traces of every experienced occurrence of words including speaker information are stored in the mental representation (Bybee, 2001; Johnson, 1997; Goldinger, 1998). These models differ in the amount of detail that is stored as a representation³.

Mental representations are believed to be located in the *mental lexicon*⁴, which is considered the language's inventory, where 'those aspects of the representation of lexical form...' are thought to be stored '...that participate directly in the process of recognizing spoken words, allowing the listener to identify the sequence of lexical items being produced by a given speaker' (Lahiri & Marslen-Wilson, 1991; p. 246).

The information that is extracted from the varying acoustic speech signal accesses the mental lexicon in order to activate its lexical entries, selecting those that are compatible with the perceived input. How these mental representations are encoded in the human brain and how detailed this information is stored in the mental lexicon remains an open issue. In the following section some of these many psycholinguistic models will be discussed.

³ 'Representation' belongs to a family of related terms that cross cognitive science. For instance, it is established to speak of a *symbolic* theory of mental representation or of brain function; written symbols such as the phoneme *b* are taken to 'symbols' in the mind. These written symbols do symbolize something, namely the entities in the mind (Jackendoff, 2003). Jackendoff (2003) proposes a model of 'cognitive structure', calling components such as the phoneme *b* 'cognitive entities' or 'structural elements'.

⁴ The mental lexicon is a central concept in word representation and is considered to be a mental store of information about words that includes semantic, syntactic information, as well as details about word forms (Gazzaniga, Ivry, & Mangun, 2002; chapter 9). Evidence is reported which suggests that modality-independent lexical knowledge is organized by categories in distinct areas of the left temporal lobe. Knowledge of words seems to be organized categorically, that is categories such as furniture, fruits, birds, etc. (Caramazza, 1996). Semantic knowledge is represented in a highly distributed fashion throughout the cortex. There is evidence that lexical/semantic access from auditory input involves the posterior lateral temporal lobe (Hickok & Poeppel, 2007).

This work will concentrate on the initial process, where the listener has to turn the time-varying speech input into some form of discrete representation in order to activate, select and access words in the lexicon. The focus will lie on an abstractionist approach - the *Featurally Underspecified Lexicon Model* (FUL; Lahiri & Reetz, 2002; 2010), which will further be discussed as it is the basis of this research and the experiments described in the experimental part of this work (chapter 3).

1.2 Models of Speech Perception

In the following section four models of speech perception will be presented and discussed. These models are considered since they all focus on the early mapping from the acoustic speech signal onto mental representations in long-term memory. These models vary considerably in how the mapping is done, how the models deal with the enormous variability in speech, in particular asymmetric assimilation processes, and more importantly how the information is stored and represented. Since the FUL model gives the theoretical background for this work it is important to discuss it in the light of other models that also deal with language representations and the early mapping mechanisms. Therefore, the other models also are presented in detail to enable a comparison with the FUL model.

1.2.1 Usage-based Model

Bybee's (2001) connectionist or network model lays out some basic statements that are applied to phonology: First, Bybee argues that experience affects representations in memory for both language production and perception. Words with high frequencies have stronger representations, which in turn can be accessed easier than low-frequency words. The lexical strength of words can undergo change as they are used in different contexts. Secondly, linguistic representations have the same properties as mental representations of other objects, assuming the brain operates in the same way in different domains. For Bybee, the processes and principles underlying language are not language specific, but pertain in general to those that govern other aspects of cognitive behavior such as memory capacity, motor control, ability to categorize experience, making inferences, etc. Mental representations are organized in categories of actual tokens of experience. Thirdly, these categories are based on similarity or identity, which organize the storage of phonological percepts (including variation). Fourthly, generalizations over different forms are not separately stored, but rather are expressed as relations among forms based on phonetic and/or semantic similarities. Fifthly, lexical organization provides generalizations and segmentation, meaning that units of features, phonemes, morphemes, or syllables surface from relations of identity and similarity that organize lexical representations.

The phonological shape of all words and frequently used phrases are stored in memory along with information about their meaning and context of use, both linguistic and non-linguistic. The storage is not a simple list, but rather a network of connections to related items that makes storage more efficient. If the lexicon is viewed as a complex network structure, a word is not necessarily an independent unit, but is instead deeply embedded in an

organizational structure. For instance, a regular inflected word as *exaggerated*, whose stem is already present in the word *exaggerate*, does not add any complexity to the lexicon and takes very little space, since it partly overlaps with other existing items. For Bybee, the question is not if *exaggerated* is indeed in the lexicon or not. Rather, whether it is there as a unit, if it has been previously used. Frequency plays a role, in that complex forms of high-frequent combinations may stand out in lexical representation.

In Bybee's model, the continuous speech stream can be chunked and associated to units of different sizes, such as features, phonemes, syllables, and rhythmic units forming schemas (patterns) at different levels. Items that occur together frequently are bound together into storage and processing units. For instance, the phrase *hit the ball* makes us group *the ball* rather than *hit the*. Once the listener has experienced which patterns occur in their language, they will be able to separate the necessary linguistic properties from unnecessary properties, and categorize them. Phonemes are also categorized as having a prototype structure, with particular tokens stored and categorized as more or less central members of a phoneme category. Even allophones⁵ that are phonetically similar are categorized as belonging to the same or related phoneme unit. For Bybee all linguistic 'objects' – phonetic sequences, words and phrases – are categorized just in the same way as non-linguistic objects such as birds, faces, dogs, etc. Central to her theory is that frequency of occurrence influences the importance of particular features that are grouped together and defined in a category. Bybee's model proposes that individual tokens of experience are stored and organized into categories without removing redundancies. Stored representations are complex and interact in complex ways, but accessing these representations is relatively direct, since the patterns found in these stored forms are represented in schemas.

How does Bybee deal with sub-phonemic variation within words and between words? She assumes that this variation is lexically specific; therefore the representations must contain sub-phonetic detail. Each word is represented with a range of variation corresponding to the actual tokens of use, experienced by the language user. Again the most frequently occurring variants will constitute the prototype of a category. Allowing frequency of use to directly affect lexical storage of items will provide a way that words are realized phonetically. Bybee for instance, argues that phenomena such as vowel nasalization are represented lexically in English although vowel nasalization is not contrastive. Phonetic tokens are classified as members of the same category if they are highly similar in their acoustic properties. Members of the same phonetic category are related across words through

⁵ An allophone is not distinctive, but rather a phonetic variant of a phoneme in a particular language or dialect; changing the allophone will not change the meaning of a word, but the result may sound non-native, or unintelligible.

lexical connections. These categorizations of tokens are all represented in memory depending on the frequency with which tokens occur or co-occur.

Furthermore, word boundaries influence phonological behavior. At word boundaries sounds come into contact with a variety of phonetic contexts, often undergoing assimilation processes. Bybee argues that these processes occur only in high-frequency phrases that in fact behave like words. Phonological alterations occur between words that are often used together, and thus become 'processing units', grouping larger chunks of language together. According to Bybee, it is within these 'chunks' that phonological changes create variant forms of words. Even phonological reductions of *I do not* such as seen in *I don't need*, *I don't have*, or reductions of *going to* to *gonna*, etc. that are repeatedly used become processing and storage units (Bybee & Schreibman, 1999; cited in Bybee, 2001).

Bybee explains in her model that phonemes are sets of phonetically similar variants (or allophones), which are clustered in groups and analyzed as contextually determined prototypes or sets of exemplars. A similar model of categorization of mental representation is the exemplar model of speech perception (Johnson, 1997; cited in Bybee, 2001; chapter 3). In this model, the fact that language listeners not only correctly identify words and phrases produced by different speakers with considerable variation, but also identify voices of different speakers, makes the language user categorize the presented variation and store all perceived tokens. A set of exemplars are represented which are associated with auditory properties and other information such as meaning or context.

A model like this requires practically unlimited memory space. It is claimed that whatever the limitations on memory, there is certainly enough memory available to allow some version of an exemplar representation (e.g. a covering map representing a space of exemplars as in Kruschke's (1992) connectionist exemplar model, where exemplars are encoded as weights rather than storing each exemplar; cited in Johnson, 1997). This model further asserts that if tokens of experience are not stored in memory, no prototype could be formed. In the case of variation, tokens that are sufficiently similar are grouped and stored together, generating a strengthened representation. This kind of model is highly dependent on the way phonemes are used in context. In this exemplar model of representation, all perceived tokens of a word are categorized and stored together with information about their contexts in which they occur. Therefore, the phonological category of a word directly represents all variation. However, Bybee argues that since language is affected by change, affecting words gradually, it appears that contextual variants of words are not stable and that reorganization of variants occurs. Therefore a simple storage of exemplars does not seem an accurate reflection of how sounds and words are structured in memory. Consequently, frequency plays a crucial role in Bybee's model, accounting for sound change and language change.

To summarize, Bybee claims that linguistic organization and storage is based on language experience, in that articulation, perception, meaning and contextual knowledge are all related to each other. Bybee argues for detailed specification of phonetic units in stored representations. These phonological representations are based on grouped and categorized units shaped by frequency and experience of language use. These linguistic precepts are organized and stored in the same way as non-linguistic precepts.

Models like this that assume all variants are stored in the lexicon have difficulties explaining the asymmetries between coronal and non-coronal sounds. According to Bybee (2001), these asymmetries can only appear through differences in frequency of occurrence. All variants are fully stored, but differences might occur in their representational strength.

1.2.2 Shortlist

The Shortlist model (Norris, 1994) assumes that a small, recurrent network generates a set of candidate words that are roughly consistent with the incoming speech input. The model relies on competition between lexical candidates, tied to specific locations in the input. Compared to other models (e.g. TRACE; McClelland & Elman, 1986) the competition takes place within a small, dynamically generated network that only considers a handful of lexical candidates at any one time, competing with each other.

The model consists of two stages: In the first stage of lexical search, a phoneme is provided as input into the recognition system and all words beginning with this phoneme are activated. Some of these words (probably those with the highest frequency) enter into the so-called 'Short-list'. With further input, the upcoming phonemes are heard and lexical search is repeated. Words that no longer match the input stream are dropped from the short-list and better matching candidates enter the short-list. The short-list therefore is continuously updated as new phonemes enter the recognition system, being seen as the generation of a small set of lexical hypothesis in a recurrent network, generating a set of lexical candidates that is based purely on bottom-up information. No top-down feedback from later processes can influence phoneme recognition or the generation of the candidate set itself.

According to Cutler and Norris (1979), phoneme identification is a race between a phonemic route and a lexical route in which the phonological representation of a word is accessed from the lexicon. The phonemic and lexical routes are completely independent and responses are determined by a first-past-the-post race. Phonemes can be identified either by a direct phonological analysis of the input (phonemic route), or by accessing the word's phonological representation from the lexicon, the representation that is available first forms the basis for the identification.

In the second stage, all activated candidate words in the short-list then compete with other (multiple) candidates for lexical selection in a constraint satisfaction network. Overlapping words inhibit each other through inhibitory links in proportion to the number of phonemes they overlap. Such a network only needs as many nodes as there are candidates in the short-list. It is important to mention that in the network, there is only an interaction between the word nodes at the lexical level, not between-levels. The direction of information flow between levels is purely bottom-up.

Since speech is a temporal signal arriving sequentially, words can begin at almost any point in the acoustic input. Shortlist solves this problem of time-invariant recognition by using a single recurrent network⁶ with time-delayed connections determined by memory span that cycle around in the mentioned short-list.

Important for this research is the form of representation assumed. The primary input of the model consists of a string of phonemes that take the form of a phonological representation. Crucially, word-representations are assumed to be explicitly stored as part of a phonological representation in the lexicon. Shortlist requires them to be sensitive to fine-grained acoustic detail during lexical activation.

The Shortlist model is based on interactive-activation networks, where each word or lexical candidate is represented by a single node assigned with an activation value. This activation value can receive either more activation, remain the same, or lose activation when it mismatches with the new input or is inhibited by a competitor. However the concept of actual *activation* values remains vague. Furthermore, activation does not directly correspond to human behavior that can be measured by speech, accuracy, or probability of response. Furthermore, the use of interactive-activation networks does not ensure that speech recognition will be optimal, especially in the case of ambiguous speech input.

⁶ This recurrent network (Elman, 1990; cited in Gaskell, Hare, & Marslen-Wilson, 1995) has a single set of input nodes corresponding to the input phonemes, and a single set of output nodes, one for each word in the network's vocabulary. Throughout the unfolding of each word, the network is trained to activate a single output node which identifies the word, and all other output nodes are then set to zero. There are time-delayed links interconnecting all units. The delayed connections in the network ensure that at each point in time, the hidden unit activation that is generated by the previous input is fed back to the hidden units. Thus, at all times, these hidden units have access to information about their state on the previous time cycle. The state of the previous cycle was self-determined by the status of the cycle before that. These delayed connections provide the network with a memory for its previous actions and enable it to integrate incoming information across time. These recurrent networks with time-delayed links between units allow generalizations to be made over temporal sequences and are valuable in processing continuous speech. To conclude, no matter when in time a word is heard, the network will be able to build up the same internal representation of the word and consequently be recognized. A network like this will recognize words at the earliest point when they become unique. Before a word becomes unique it will activate all members of the cohort.

For just these reasons, Norris and McQueen (2008) felt the need to replace the interactive-activation network of Shortlist where the output is a pattern of word activations over time and developed **Shortlist B**. Here, the concept of activation is replaced by the concept of likelihood and probability based on Bayesian computations. In Shortlist B the output is a list of probabilities of words. It still shares most of the key assumptions of Shortlist (A), for example the nature of pre-lexical and lexical representations and bottom-up flow of information, but with the benefit of making Bayesian decisions.

Since speech does not consist of strings of unambiguous isolated words, but rather is phonetically indefinite and unfolds over time in a continuous stream, Shortlist B uses Bayesian inference. These allow word recognition to deal with ambiguity by combining the incoming speech input with knowledge of the prior probabilities of sounds and words. Bayes's theorem allows one to calculate the conditional probability of each word given the available input, thereby being able to specify and update new incoming information. Once these conditional probabilities have been computed, decisions can be made either by selecting the most probable sound and word at a given time or by making a response.

In Shortlist B, words do not have activation values as stated above, and there is no direct inhibition between word candidates, but rather the search of the best matching input is done by performing computations on path probabilities. Path – in the sense used in Shortlist B – refers to strings of phonemes. However, the input of the word-recognition process is no longer a string of phonemes, since acoustic information needs to be modulated on a much finer time-scale. The input of the new model is a sequence of multiple phoneme probabilities over three time slices per segment. In particular, the input signal might have been generated by more than one phoneme, being more ambiguous in the input due to e.g. co-articulation or environmental noise. Therefore it is possible to estimate phoneme probabilities in terms of likelihood functions. One key assumption of the model is that listeners acquire knowledge of phoneme likelihoods early in life, which are compared against the incoming phoneme probabilities.

The word recognition in Shortlist B proceeds as follows: A given speech input with phoneme probabilities results in path probabilities which are also sequences of words that make up probabilities of words. These path probabilities are compared to lexical entries that have been learnt and estimated from the probability of each word given the path of phoneme-probabilities in the input. These probabilities are continuously updated. The Bayesian computations assure that low probability words (and paths) automatically fall out of the competition list, so that only the short-list of best candidates influences word recognition. The probability paths that a word lies on also influence overlapping candidates. They will only compete to the extent that they lie on paths with high probability.

The model makes it possible to calculate the relevant probabilities and estimate actual reaction time (RT) and error rates during lexical decision tasks and phoneme categorization. Furthermore, Shortlist B makes it possible to estimate word frequency, lexical neighborhood frequency and density and deal with segmentation and mispronunciations in the continuous speech stream – how the likelihood of different realizations of words can modulate word recognition. Even asymmetries in speech perception can be explained by likelihood functions, for instance that one phoneme might be misperceived more often than others.

To sum up, the feature that sets apart Shortlist B from its predecessor is the modeling of the matching process being determined by likelihood functions, rather than purely based on similarity. It no longer uses a serial search of the lexicon; only words that are real potential candidates are considered.

Despite being very different models, Shortlist B and (A) (Norris, 1994, Norris & McQueen, 2008) and the Usage-based model (Bybee, 2001) share the view that the core representation of a lexical form is fully specified and stored in the mental representation. Further, they both assume that the stored representation is fully specified in its citation form of the incoming speech signal. Phonetic input that derives from this citation form – as is the case for example in assimilation processes – is accepted as tokens of the form in memory. These two models differ, however, in the way they deal with this variation. Where Shortlist B computes and updates phoneme probabilities and path probabilities that make up probabilities of words, comparing them to lexical entries where the best candidates influence word recognition, Bybee's model relies on experience, where language input is grouped together and build categorized units that are then recognized.

In the following I will describe a model dealing specifically with phonological variation of given word forms that account for the fact that words can be phonetically realized in many different ways.

1.2.3 Distributed Model of Speech Perception

Gaskell and his colleague's *Distributed Model of Speech Perception* (Gaskell, Hare, & Marslen-Wilson, 1995; Gaskell & Marslen-Wilson, 1997, 1999; Gaskell, 2003) tries to model the process of speech perception as a direct mapping from low-level featural information to distributed representations of lexical knowledge and form. They make the following assumptions: Lexical knowledge is represented in a fully distributed fashion (connectionist model). Different forms of lexical knowledge (e.g. phonology, semantics) are represented in parallel and accessed simultaneously. Speech input maps directly and

continuously onto lexical knowledge and finally the lexical access process operates under maximum efficiency. The model focuses on the *retrieval* and *access* of lexical phonological and semantic information, thus the mapping process itself, not so much on the structure of lexical phonological representations.

The model is described as a connectionist recurrent network (see also footnote⁶; Elman, 1990; Norris, 1990; cited in Gaskell et al., 1995) in which patterns of phonetic features spread across time (Figure 1). Phonological features are mapped directly and continuously onto lexical representations. The featural input passes through a set of hidden units that have access to the hidden units of the previous time phase through recurrent links to allow the alignment of important information as well as for generalizations to be made over temporal sequences of patterns. These hidden units are also connected to output units, representing the phonology and semantics of words that are held in the speech stream. The input into the model is a representation of the speech signal, idealized as a set of features that are flexible and un-integrated in order to accommodate some sub-phonemic variation such as assimilation processes. The output nodes can be semantic or phonological.

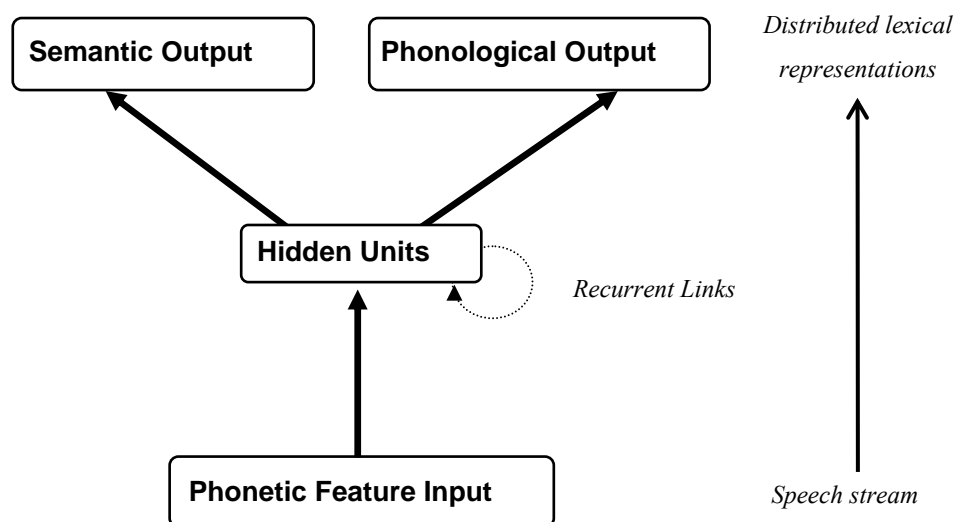


Figure 1: A distributed model of speech perception (Adapted from Gaskell & Marslen-Wilson, 1997, 1999).

The semantic output node is a distributed representation of the stored semantic knowledge about words (non-phonological knowledge about a word). The phonological output node encodes a representation of the underlying phonological structure of words (see Figure 1). This phonological knowledge is capable of making phonological judgments on

words and word-like nonwords, where context-dependent phonological inference processes assess the phonological variability of surface changes due to assimilation. Judgments are made about the form of words and are a product of the mapping process to semantic knowledge. These form representations are considered highly abstract, in the sense that they only contain the invariant properties of a word. The unassimilated most frequent form of a word is identified as the underlying form (Gaskell et al., 1995).

Lexical representations are seen as distributed patterns including its phonological, morphological, semantic and syntactic specifications. All types of information deriving from the speech input signal is treated as output of the system, with detail preserved in the mapping onto these forms of knowledge. This means as speech information enters the recurrent network the activation of words is reflected by a ‘blend’ of their representations. Once a word onset is presented, the network produces a blend of representations of all matching words that match this onset. The identity of words often is ambiguous early on in the word, where the network produces ‘cohort-like’ behavior⁷. The relevant information about a word is extracted and if more than one lexical entry matches the speech input, the output of the model should reflect this ambiguity and activate the stored knowledge about these candidates. As more speech comes in, this blend is purified to represent the reduced set of words that still match the input. This process continues until just one word matches the input. These recurrent networks develop as a reflection of experience that allows the mapping mechanism.

The network must consider multiple hypotheses in parallel about the lexical identity of incoming speech which are represented by the phonological and semantic output nodes. This assumption is entertained by many current models of speech perception; however unique to this distributed model is that speech is mapped directly onto distributed representations of lexical knowledge. Thus, multiple lexical candidates are evaluated and handled as hypotheses on this one level of representation and not on some other independent competition level (see Shortlist, Norris, 1994). These hypotheses are represented in different ways for two output nodes: For the phonological output the network simply activates the phonemic nodes (assuming the system can identify the onset of a word) that correspond to the segments presented so far. For example, when the onset of /hænd/ is heard words such as *hand*, *ham* or *had* are activated as their representations reflect similarities in the phonological

⁷ ‘Cohort-like’ behavior goes back to the Cohort model (Marslen-Wilson, & Welch, 1978; Gaskell & Marslen-Wilson, 1999) where the processing of speech starts with the very first sound that a listener has identified as the onset of a word. Initially, when not all perceptual information is available, more than one representation will be activated as more than one representation will fit the first part of the input. Many words are then simultaneously activated, causing a blend of the relevant representations (word initial cohort). As more perceptual information becomes available the number of activated representations narrows down leading to a reduced set of words that still match the incoming signal. This process goes on until only one form remains and the word can be recognized.

form. As more input is perceived, the blending will be reduced as only the matching words stay active. This process continues until one clear representation remains.

The output at the semantic level seems more complex due to the random relationship between form and meaning of words (Gaskell & Marslen-Wilson, 1997).

Other models such as Shortlist (Norris, 1994, Norris & McQueen, 2008) have taken a more serial approach to the representation of different types of information. The selection of word candidates is seen as a parallel process of competition (via lateral inhibition). This distributed model operates on a single distributed level of representation, where speech is directly mapped onto lexical representations of form and meaning. Activation is encoded implicitly on the similarity between the network's output and each word's distributed representation.

The model treats ambiguity in the speech signal by activating a 'blend' of the relevant distributed representations. In the course of hearing spoken words, a blend of matching lexical representations is built up and is continuously modified through new incoming information. This blending process takes place for both phonological and semantic nodes. The resulting activation pattern is a weighted average of all relevant elements. The simple recurrent network allows the model to pick up statistical information reflecting learned probabilities biasing its output towards more frequent word candidates. This means that these blended representations are weighted in terms of frequency of the matching candidates that are learned and based on statistical properties of the speech stream.

But how does the mapping from phonetically variant surface forms to their underlying representations actually work in this model? It is proposed that connectionist-learning principles are applied to this mapping. Assimilation can be governed by phonological inference, using the context of phonological changes to infer the underlying representation. The speech system learns to apply rules of inference through statistical experience. Phonological inference is therefore a contextual sensitivity in the processing of surface segments. It analyses segments with reference to their phonological context in order to identify the underlying stored segment of the original form. Thus, the /nd/ underlying the surface [m] in [hæmbæg] (underlyingly *hand bag*) can be inferred from the place of the following labial [b], since this assimilation occurs frequently in the varying speech stream. Lexical access will be successful if the phonological inference mechanism analyzes the variant speech so that the underlying fully specified form representation can be recovered and identified in the output node.

Evidence for this inference process comes from studies examining the effects of phonological assimilation of place of articulation on lexical access (Gaskell & Marslen-Wilson, 1996, 1998, 2001; Gaskell & Snoeren, 2008). Assimilated items in phonologically viable contexts produced good facilitation, whereas the same alterations in unviable contexts showed no significant activation of the critical word. These results are explained in terms of a context-sensitive inference process that resolves ambiguity during speech perception providing an additional prerequisite in mapping ambiguous speech into lexical representations.

To summarize, variable surface forms of speech are extracted in forms of features that are mapped onto distributed and fully specified representations of lexical form. Inference processes based on statistical properties and experience solve the problem of variation. The model uses a simple recurrent network that maps the surface form (which might be assimilated or distorted) onto underlying stored representations. The network compensates for variation once the following context has been identified as possible for variation to occur.

In sum, we have examined three models of speech perception. These all incorporate the view that the core representation of lexical form is a fully stored representation in the mental lexicon. They deviate in the amount of detail stored and thereby also deal with language variation differently. As previously discussed, the first model (*Usage-based model*; Bybee, 2001) assumes that lexical entries are made up of very detailed memories derived from grouped and categorized units based on the experience by the native speaker (Bybee, 2001; Johnson, 1997). These memories are continuously updated and comprise all surface variants of a given word.

The second model (*Shortlist (A) and B*; Norris, 1994; Norris & McQueen, 2008) presumes that the speech process evolves on a phoneme-by-phoneme basis or by phoneme likelihoods with fully stored phonemic representations. Phonemes are mapped onto a pre-lexical phonemic level, and are compared against whole word forms specified in the lexicon. The word search for the best matching input is done by performing Bayesian computations on path (string of phonemes) probabilities. These path probabilities being continuously updated are compared to lexical entries that have been estimated from the probability of each word given the path of phoneme-probabilities in the input. This model compensates for variation by modeling the likelihood of different realizations of sounds and words. The matching process is determined by likelihood functions, rather than purely based on similarity (Norris & McQueen, 2008).

The third model (*Distributed Model of Speech Perception*; Gaskell et al., 1995; Gaskell & Marslen-Wilson, 1997, 1999) assumes less detailed and more abstract lexical

representations. A recurrent network maps the surface form directly onto underlying stored representations (e.g. lexical knowledge such as phonology or semantics), which are distributed and accessed simultaneously. It highlights the active process of compensation for the variability in speech, rather than the organization through detailed representations. Segments are analyzed in the context of phonological inference rules that specify the phonological contexts in which different variants can take place (Gaskell & Snoeren, 2008; Gaskell & Marslen-Wilson, 1996, 1998, 2001). These rules dictate that the interpretation of the variant surface form requires the appropriate phonological context in order to map these variable surface forms onto fully specified representations of lexical form.

Furthermore, all three models assume that experience with a language – for instance, learning that certain words occur more frequent than others, and that some sounds change in certain contexts – play a major role in accessing words and perceiving language.

The following and fourth model diverges from the previous three in its fundamentally different assumptions about the mental representation of lexical form. Lexical representations are assumed to be highly abstract, where only distinctive, non-redundant information is specified in the mental lexicon. This *Featurally Underspecified Lexicon (FUL) model* (Lahiri & Reetz, 2002, 2010) can deal with variation through sparse lexical representations thereby being insensitive to surface variation in speech. It tries to explain – more than the other models – how language is represented and processed in the human brain by looking at the type of lexical representations and having basic assumptions on how they are accessed in the brain. This model forms the basis of this research and the predictions of all the experiments here are derived from this model's assumptions.

1.2.4 Featurally Underspecified Lexicon Model (FUL)

This recognition model of feature representation (Lahiri, 1999, 2000; 2007; Lahiri & Reetz, 2002; 2010; for review see also Fitzpatrick & Wheeldon, 2000) evolves around both language comprehension and production with a strong focus on speech perception. In this representational account, the acoustic speech stream is converted into distinctive phonological features and directly mapped onto stored lexical representations activating word candidates. All features extracted from the speech signal are compared to their corresponding underlying representations by means of a ternary matching system involving *match*, *no-mismatch*, and *mismatch* (later in this section this matching algorithm will be explained and discussed in more detail). There is no intermediate conversion from features to segments. All word candidates, which do not mismatch with their initial feature set from the acoustic signal, are activated, together with their phonological, morphological, syntactic and semantic information defining a lexical entry. This parsing begins as soon as features are

extracted from the signal. The lexicon works in tandem with grammar to conceive new candidate representations or reject a present candidate whenever new information becomes available.

Furthermore, the model assumes that each morpheme has a single, unique underlying lexical representation. No phonological, morpho-phonological or post-lexical variants are stored. For the low-level process of comparing the incoming signal with the stored information in the lexicon, only phonological representations are considered in this model. These phonological representations are ‘abstract’ in the sense that they abstract away from variability of the acoustic speech stream. Furthermore, they are organized in hierarchically structured features allowing relationships between different features grouping them together as functional units. Importantly, not all features are stored in the mental representation, only distinctive and contrastive features are stored and specified. Predictable information is not stored and thus *underspecified*. The precise specification depends on universal and language-specific requirements.

As already mentioned, the lexical representation in this model consists of highly abstract features that make up words or rather morphemes with no variants stored and with only one single underlying representation available. Phonological features are abstract entities and certain features group together. Many proposals on the hierarchical organisation of features have been made, usually characterized as a feature tree (Lahiri, 2000) represented with root nodes where features are hierarchically organized into functionally related groups (such as place- or manner-of-articulation features).

Distinctive features and their organization have been widely discussed and are central to phonology (Jakobson, Fant, & Halle, 1952; Chomsky & Halle, 1968; Sagey, 1986; Clements, 1989; Clements & Hume, 1995; Halle 1995; Halle & Stevens, 1991; cited in Lahiri & Reetz, 2010). The FUL model organizes these features in the feature tree as seen in Figure 2. The features and their organization are based on universal principles of phonological variation that can express segmental contrasts of all languages of the world. The structure of the features is dominated by the root node grouping together smaller feature nodes and ending in the terminal distinctive features. The root node forms two binary features such as [CONSONANTAL] or [VOCALIC] and [SONORANT] or [OBSTRUENT], which are the major feature classes of all languages. It still is discussed in the literature whether consonants and vowels share the same features and if they can be organized under the same root nodes (e.g. Chomsky & Halle, 1968; Sagey, 1986; Halle, Vaux & Wolfe, 2000; Clements, 1989; Lahiri & Evers, 1991; Lahiri & Reetz, 2002; cited in Lahiri & Reetz, 2010). In the FUL model consonants and vowels share the same place features. A sub-class of features – manner of articulation – is assumed to be universally dependant on the root node,

in that, for instance, [NASAL] is [SONORANT]. On the other hand, [STRIDENT] and the related feature of [CONSTRUCTION] such as [PLOSIVE] is [OBSTRUENT].

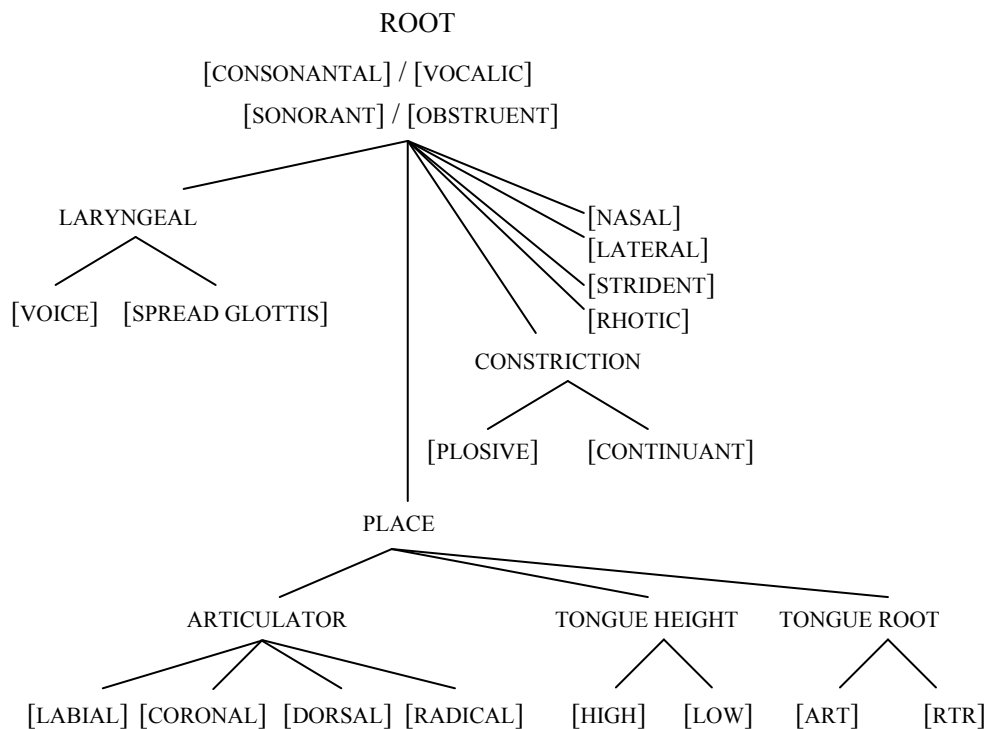


Figure 2: Feature organization of the FUL model (Adapted from Lahiri & Reetz, 2010).

The sub-class of place-of-articulation features is split into three independent nodes, an ARTICULATOR node, TONGUE HEIGHT, and TONGUE ROOT, which are identical for vowels and consonants. These phonological features are only stored in the underlying representation when they are unpredictable, all other predictable and redundant features – deriving from rules – are underspecified. The feature [CORONAL] under the ARTICULATOR node for instance is assumed to be (almost) universally underspecified in the mental lexicon, other features may also be underspecified depending on the feature assignment to establish a phonemic contrast in a certain language. There have been different reasons for assuming that [CORONAL] is underspecified (Lahiri, 2000). Coronals are more frequent in speech and are special in the sense that within the coronals several places of articulation can contrast phonemically (e.g. dental, palatoalveolar, palatal, retroflex, etc.) within one language. They appear to be more tolerant allowing more combinations than other places of articulation. Additionally, they appear to be the default place in many languages (Paradis & Prunet, 1991). Finally, coronals are more likely to assimilate to other places of articulation than the

other way around; hence, their ‘placelessness’ makes them more vulnerable to acquiring the place of neighboring segments.

The other place features [LABIAL] and [DORSAL] are contrarily assumed to be specified for place. Further, it is assumed that features such as [NASAL], [STRIDENT] and [LATERAL] are specified for manner of articulation if a feature contrast is established. [ROTHIC] and [PLOSIVE] however are supposed to be underspecified for manner of articulation. Other features depend on language-specific phonological systems⁸. The feature [NASAL] in vowels is considered to be specified only when it contrasts with oral vowels. For instance, English only has oral vowels. In cases of consonant-vowel-consonant (CVC) or consonant-vowel-nasal consonant (CVN) clusters, the vowel is always underlyingly oral, despite of variation where the vowel will be nasalized when followed by a nasal consonant. This will lead to surface variants such as *hand* [hænd], or *ban* [bænd]. This surface nasalization is predictable and therefore does not need to be specified in the lexicon as feature [NASAL], as it only occurs when the nasal feature spreads to the preceding vowel. This is different for Bengali, where both oral and nasal vowels are underlyingly represented since they are distinguishing as in the example of [pāk] (‘*slime*’) and [pak] (‘*cooking*’), where the vowel in the CVC-sequence is specified as underlying [NASAL]. Thus, across both languages the surface nasal vowel can come from two different sources: from an underlying nasal vowel in the case of Bengali or from regressive assimilation, which can be the case both in Bengali and English (Lahiri & Marslen-Wilson, 1991, 1992).

The mapping of features to lexical representation in the FUL model depends on a ternary matching logic, where a *match* occurs if both the acoustic signal and the mental representation have the same features. A *mismatch* occurs if the feature extracted from the signal contradicts the feature that is stored in the lexicon. In the case of a *match*, the lexical word candidate will be activated or increased in its activation, in the case of a *mismatch* a word will be excluded from the list of possible candidates. A *no-mismatch* appears if either no feature is extracted from the signal although there are features stored in the representation (e.g. in the case of reductions, or in the case of vowel height where the feature [MID] is not assumed to be extracted from the speech signal) or if a feature is extracted from the signal but does not find a corresponding feature in the lexicon. This is the case for lexically underspecified features. Crucially, based on this assumption, there is an asymmetry in the matching from the signal to the lexicon. No-mismatch situations do not lead to the rejection of candidates, but receive less activation than in the case of a match between the signal and

⁸ Eliminating predictable and redundant information from the mental lexicon is phonologically motivated. For instance, not only features, but also suprasegmental information (e.g. tone in non-tonal language) that are not distinctive in a given language will be underspecified (Fitzpatrick & Wheeldon, 2000).

representation. The signal simply does not contradict a candidate. Only mismatching features lead to the rejection of candidates. The level of activation is measured on the basis of a scoring formula:

$$\text{Score} = \frac{[\text{Number of Matching Features}]^2}{[\text{Number of Features in the Lexicon}] \times [\text{Number of Features from the Signal}]}$$

The scoring of these segments (feature bundles) gives a word score and with it a ranking in the list of possible candidates.

This matching logic together with the underspecification of certain features can account for asymmetries found in assimilation processes. This can be demonstrated in the following example: An incoming speech sound can vary in terms of place-of-articulation features extracting either [LABIAL], [CORONAL] or [DORSAL]. As mentioned before, according to FUL the feature [CORONAL] is underspecified in the mental representation, while the features [LABIAL], and [DORSAL] are assumed to be specified. If for instance, a dorsal plosive such as [g]⁹ is perceived in the acoustic signal the feature [DORSAL] will be extracted and directly mapped onto its lexical representation. It *matches* with the underlying representation of a /g/¹⁰ being specified as [DORSAL]. It *mismatches* with a representation of a [b], which is specified for [LABIAL] place of articulation and therefore will conflict. Consequently, hearing an initial [g] will activate word cohorts starting with a [DORSAL], but will reject all candidates beginning with a [LABIAL]. If the [g] is mapped onto a coronal /d/ which is underspecified for place in the representation, a *no-mismatch* occurs. The feature [DORSAL] is mapped onto an ‘empty space’ in the underlying representation for place of articulation. Hence, the dorsal sound [g] can activate lexical representations of /g/ as well as /d/, but /d/ to a smaller degree by the means of the scoring formula. In the reversed case, when a [d] is perceived in the speech stream, its [CORONAL] place will be extracted from the signal and will *mismatch* the representation of /b/ and /g/ which are both specified once for [LABIAL] and once for [DORSAL] creating a conflict in both cases. This leads to a *mismatch* in lexical activation, since coronal sounds will only activate lexical entries with other coronals (i.e. underspecified segments), but labial or dorsal sounds can activate their own representations as well as coronals.

The ternary matching logic makes it possible for the perceptual system to cope with speech variation without having to store all details in the lexicon. For instance, place assimilation can be well explained with this logic. Place assimilation causes surface variation


⁹ Sounds between square brackets as [x] will always represent the surface phonetic form of the acoustic incoming sound.

¹⁰ Sounds between slashes as in /x/ always mark the underlying representation of a sound stored in the mental lexicon.

such as **valib* as in *valid parking* where words actually change into nonwords. In this example the incoming sound [b] of **valib* extracts the feature [LABIAL] and maps onto the underlying representation of the [CORONAL] /d/ in *valid* which is underspecified. In this case **valib* as variant will activate its original form *valid*. This is not the case if features are reversed: nonwords such as **biraffe* or **diraffe* will not activate the original word *giraffe*. The [b] or [d] will extract the features [LABIAL] or [CORONAL] from the signal and map onto the [DORSAL] feature of [*g*]iraffe in the underlying representation. This causes a mismatch or a conflict with the specified [DORSAL] underlying representation and will lead to the rejection of the nonwords in the lexicon.

These asymmetrical activation patterns occurring, for example due to assimilation processes can be explained by assuming underspecified representations in the lexicon. Successful lexical access predominantly depends on the degree of match between specified features of an underspecified lexical representation and the corresponding speech input. Thereby, the system over-generates possible word candidates but does not include impossible candidates. A further example of the FUL logic and its ternary matching system is given in Table 1.

Table 1: The ternary matching logic of the FUL model, which compares the incoming speech signal against the representation in the lexicon in order to activate word candidates:

Features extracted from the speech signal at one instance in time		[CONS]	[NASAL]	[LABIAL]	
	→	[CONS]	[NASAL]	[LABIAL]	
Features stored in mental representation		↓	↓	↓	
/m/		[CONS]	[NASAL]	[LABIAL]	
		match	match	match	→ candidate (higher score)
		↓	↓	↓	
/n/		[CONS]	[NASAL]	[]	
		match	match	no-mismatch	→ candidate (lower score)
		↓	↓	↓	
/z/		[CONS]	[STRIDENT]	[]	
		match	mismatch 'conflict'	no-mismatch	→ rejected

The distinctive features in the acoustic signal are compared to those in the mental representation. Through the ternary matching logic of – match – no-mismatch – mismatch – word candidates are activated if the feature does not mismatch with the feature computed from the signal. This process can handle certain variation in speech. [FEATURE] = feature is stored in the lexicon; [] = feature is underspecified, [CONS] = [CONSONANTAL] (Adapted from Lahiri, 1999)

To summarize, the FUL model with its assumption of the underspecification of contrastive feature representations and the ternary mapping process offers a good solution

for dealing with asymmetries in assimilation and speech variation in perception and production. Variation here is solved at a representational level where monovalent and abstract features are stored and organized in a hierarchically structured manner. Features are grouped onto tiers allowing relationships between them. The mapping process from signal to the lexicon crucially depends on the absence or presence of these features in the representations in the lexicon.

For this line of research, the features of place of articulation as well as manner of articulation will be important, in particular the place features [LABIAL], [CORONAL], and [DORSAL], as well as the manner features [NASAL], [STRIDENT], and [PLOSIVE].

The assumptions made by FUL contradict models which assume full specification of lexical representations, and that deal with variation either by storing all variant forms (Bybee, 2001; Johnson, 1997) or by compensating variation at a surface level – by modifying the mechanisms that map from speech to lexicon – using context information and experience (Gaskell & Marslen-Wilson, 1996, 1997, 1998, 1999; Norris, 1994; Norris & McQueen, 2008).

The hypotheses that the FUL model makes, however, are quite precise and testable which has already been done in a variety of behavioral and neurophysiological studies testing different features. Most studies examined place-of-articulation features, in particular on features of the ARTICULATOR node, such as place features (Lahiri & Marslen-Wilson, 1991; Ghini, 2001; Lahiri & Reetz, 2002, 2010; Obleser, Lahiri & Eulitz, 2003, 2004; Eulitz & Lahiri, 2004; Wheeldon & Waksler, 2004; Walter & Haquard, 2004; Gumnior, Zwitserlood, & Bölte, 2005; Friedrich, Eulitz, & Lahiri, 2006; Friedrich, Lahiri, & Eulitz, 2008; Scharinger, Reetz, & Lahiri, 2009; Scharinger, 2009; Scharinger & Lahiri, 2010; Zimmerer, Reetz, & Lahiri, 2009; Cornell, Lahiri, & Eulitz, 2008, 2009), others concerned height features (Felder, 2006; Felder, Friedrich, Lahiri, & Eulitz, 2008; Scharinger et al., 2009; Scharinger, 2009). Further work has also been done on speech production and perception of reduced speech (Zimmerer et al., 2009; Zimmerer, 2009). Some of these studies will be discussed in further detail in the following section in particular those which are relevant to this work.

Manner-of-articulation features have been looked at in vowels cross-linguistically (Lahiri & Marslen-Wilson, 1991; 1992). For the very first time different manner-of-articulation features with consonants will be looked at in this line of research (Cornell, Lahiri, & Eulitz, submitted).

The next chapter will introduce some of these studies as evidence for the assumptions that are proposed by FUL.

1.3 Underspecified Representations

1.3.1 Evidence from Psycholinguistic Behavioral Experiments

Lahiri and Reetz (2002) investigated the assumptions made by the FUL model. The underspecification of the feature [CORONAL] in word final as well as word medial consonants was tested in a cross-modal semantic priming experiment with a lexical decision task. In their first experiment the conditions contrasted manipulations of the place feature for underspecified or specified segments using final nasal consonants. In a semantically related condition subjects were auditorily presented with German prime words such as *Bahn* ‘railway’ or *Lärm* ‘noise’. Following the prime a semantically related target word (e.g. *ZUG* ‘railway’, *KRACH* ‘noise’) appeared on the screen. In the test condition, the place of articulation of the final segment of the primes was altered into pseudowords in that a coronal segment (e.g. *Bahn*) changed to a noncoronal segment (e.g. **Bahm*), and a noncoronal final segment (e.g. *Lärm*) turned into a coronal final segment (e.g. **Lärn*). A control condition used semantically unrelated primes such as *Maus* ‘mouse’ or *Blatt* ‘leaf’ for the target *ZUG* or *KRACH*. The results showed that *Bahn* primed the semantically related word *Zug*, and importantly as assumed by the underspecification approach, the pseudoword **Bahm* also primed the target *ZUG*, compared to the control condition. The results for the noncoronal prime words were different: *Lärm* significantly primed its related target *KRACH*, however manipulating its place feature to a coronal did not result in priming, so **Lärn* did not prime *KRACH*, not differing from the control condition. These results show that lexical processing is not disrupted when noncoronals as in **Bahm* are extracted from the speech signal and mapped onto the lexical representation of an underspecified [CORONAL] creating a no-mismatch situation. Contrary, words that are underlyingly specified for a [LABIAL] place such as in *Lärm* were not activated by mismatching coronal segments in the speech signal as was the case for **Lärn*.

Comparable results were obtained for word medial nasal consonants using the same experimental setup. Coronal prime words like *Düne* ‘dune’, as well as their changed pseudowords **Düme* activated the semantically related target *SAND* ‘sand’. However, only noncoronal primes (e.g. *Schramme* ‘a scratch’) could prime their semantically related target (e.g. *Kratzer* ‘a scrape’), but the changed pseudoword prime **Schrann*e could not activate the target. These results support the underspecification in phonological processing, showing an asymmetry in activation that fully specified or purely experience-based representations could not account for.

Independent evidence for the underspecification account was found by **Wheeldon and Waksler (2004)**. They tested the underspecification of the feature [CORONAL] of word

final consonants in lexical representations and the nature of the online mapping process in a cross-modal repetition priming experiment. Cross-modal repetition priming is assumed to reflect the activation of lexical representations, and response times indicate the degree to which the lexical entry of the target is accessed. Each prime word was embedded in sentences. Four versions of each sentence were constructed, so that the prime was either followed by a context word licensing an assimilation of coronal consonants, or not licensing it. The presented sentences contained a prime ending in either a coronal (e.g. *wicked*) or a noncoronal (e.g. *frantic*) consonant. At the offset of the critical prime word the visual target (e.g. coronal target: *WICKED* or noncoronal target: *FRANTIC*) appeared and subjects had to make a lexical decision response. In an unchanged condition prime words were presented in their unchanged original forms (*'They heard there was a wicked prince in the castle'* / *'She never had frantic moments with the twins'*). In a change condition, the place of articulation was altered either to match the following context in the appropriate condition (e.g. *'They heard there was a *wickib prince in the castle'* / *'She never had *frantip moments with the twins'*.) or to mismatch with the following context in an inappropriate condition (e.g. *'They heard there was a *wickib goast in the castle'* / *'She never had *frantip days with the twins'*). The control condition was phonologically and semantically unrelated to the prime words (e.g. *'She never had idle days with the twins'*). This experiment used identical context sentences for comparable conditions. Significant priming was found in comparison to the control condition for words ending in a coronal consonant (underspecified), with no difference between the unchanged original form (e.g. *wicked*) and the changed prime word (e.g. *wickib*). The appropriateness of the following context did not have an influence as an underspecification approach would predict. The noncoronal (specified) consonants however showed a different pattern of results. Unchanged original prime words (e.g. *frantic*) were responded to faster than changed primes (e.g. **frantip*), both being faster than the control condition.

Neither segment change, nor appropriateness of context modulated the degree of priming observed in this study that supports underspecification of the feature [CORONAL] as well as a context-independent mapping mechanism.

These two studies by Lahiri and Reetz (2002) and Wheeldon and Waskler (2004) concentrated on place-of-articulation contrasts.

However there is further evidence for manner of articulation for vowels as well. It was argued that the feature [NASAL] in vowels is considered to be specified only when it contrasts oral vowels. **Lahiri and Marslen-Wilson (1991, 1992)** tested this kind of lexical processing in lexical entries under conditions where the same surface feature [NASAL] varied in its underlying representation cross-linguistically. Three types of oral and phonetically

nasal vowels were presented in a gating task (Grosjean, 1980) to Bengali and English subjects. The first group consisted of simple CVCs (e.g. [kap]) and the second group of CVN words ending in a nasal consonant with an assimilated nasalized vowel (e.g. [kæ̃n]). The Bengali group was additionally presented with CVC words containing an underlying nasal followed by an oral consonant (e.g. [kãp] ‘slime’). The stimuli were presented in gradually increasing gates where the subjects were asked to name the word they thought they were hearing. The responses before the final consonant were of interest during vowel perception and were analyzed. Hearing a nasal vowel, English speakers used the vowel nasalization to predict CVN words, as a cue for the following nasal consonant. However Bengali subjects reported a mixture of oral and nasal words with a strong bias towards CVC words rather than CVN words. They interpreted the surface nasality as an underlying nasal, accessing therefore more CVC words. The nasalized vowels completely matched the lexical representation. This indicates that nasal vowels are specified in the mental lexicon for manner of articulation in Bengali, but not in English as predicted by an abstract underspecification account.

1.3.2 *Electrophysiological Evidence*

A number of studies have used neurophysiological techniques such as magnetencephalography (MEG) and electroencephalography (EEG) in particular event-related brain potentials (ERP) to study sound representation and possible differences (among them) in representation. The advantages of these kinds of measures are that they occur as a correlate of behavior – similar to the behavioral measures of button-press – and a direct online reflection of brain activity. This can be used to examine the functional organization of the brain for language and language processing. Importantly, these measures are as close to immediate and online now as is currently technically possible. The ERP waveform provides information about very early sensory and perceptual processes, which is of great interest to this line of research. Further details about the EEG and ERP will be given in chapter 1.4 ‘Methodology: Possibilities of Studying the Mapping from Sound to Representation’.

Underspecification of the feature [CORONAL] is assumed to be a universal principle of the mental lexicon and not dependent on its position in a word. This assumption has been validated by two EEG studies by Friedrich and colleagues (2006, 2008). Their first experiment examined degrees of abstractness word medially. **Friedrich, Eulitz, and Lahiri (2006)** used an auditory lexical-decision task in an EEG experiment examining N400 ERP components, which tap into lexico-semantic processes. They used German words with coronal (e.g. *Horde* ‘horde’) or noncoronal medial consonants (e.g. *Probe* ‘test’). These

words were altered in place of articulation of the medial consonant turning them into pseudowords so that words with a medial coronal become a pseudoword with a medial noncoronal (e.g. *Horde* → **Horbe*) or reversed, words with medial noncoronals became a pseudoword with a medial coronal (e.g. *Probe* → **Prode*). According to the theory of underspecification, word medial coronals are underspecified for place; therefore, noncoronal variants should not mismatch with this empty place slot in the underlying representation and are expected to activate words like *Horde*. This mapping should not occur in the reversed case such that altered coronal pseudowords like **Prode* should not activate its real word *Probe* since it mismatches with its underlying feature [LABIAL]. The behavioral results revealed pronounced error rates for the noncoronal pseudowords like **Horbe* compared to the coronal pseudowords like **Prode*. It appeared to be easier for subjects to recognize **Prode* as pseudoword and more difficult for them to reject **Horbe* as pseudoword since it activated its lexical entry of *Horde* thereby producing more errors. The N400 results showed asymmetrical effects as well: the amplitudes for coronal pseudowords (e.g. **Prode*) started to differ about 150 ms earlier from their noncoronal base words (e.g. *Probe*) than the noncoronal pseudowords (e.g. **Horbe*) from their coronal base words (e.g. *Horde*). The noncoronal pseudowords were accepted as variants activating their original coronal base word, whereas the coronal pseudowords were rejected immediately causing a pronounced and earlier N400 effect, as well as lower error rates. These results support the idea of underspecified lexical representations and an asymmetric mapping of phonological features onto mental representations in word medial positions.

Additionally **Friedrich, Lahiri and Eulitz (2008)** examined whether initial segments of coronal and noncoronal pseudoword variants activated lexical entries of words in two further ERP studies. They used a lexical decision task with and without fragment priming. In the fragment priming experiment half of the target words began with a noncoronal (e.g. *GRENZE* ‘boarder’) and the other half with a coronal segment (e.g. *DDRACHEN* ‘dragon’). These words were primed with fragments of their first syllable such as *Drach-* or *Gren-*, as well as with fragments where the place of their initial consonant was reversed to *Brach-* or *Dren-* eventually leading into pseudowords like **Brachen* or **Drenze*. The word targets were primed by their original fragments and their changed pseudoword fragments. The event-related potentials indicated that noncoronal pseudoword fragments as in **Brach-* activated their coronal base word *Drachen*, while coronal pseudoword fragments as in **Dren-* did not activate their noncoronal base word *Grenze* as effectively. These results indicate that certain word onset variations do not hinder lexical activation.

In their lexical decision task *without* fragment priming Friedrich et al. (2008) used the same stimuli as described above but with whole words. In this experiment the same

results were found: non-coronal pseudowords activated coronal words, whereas coronal pseudowords did not activate non-coronal words. For instance pseudowords like **Brachen* activated *Drachen*, but **Drenze* did not activate *Grenze*. This neurophysiological evidence supports the assumption of the underspecification of the [CORONAL] place of articulation. Models assuming fully specified information where all variants are stored or context sensitive approaches could not account for the asymmetries found in these experiments.

Until now these studies focused on the activation of lexical items in the mental lexicon under attentional processing. The following experiments rather deal with the nature of stored representations thereby using pre-attentive measures.

The nature of mental representations of phonological features has also been studied in isolated vowels (Eulitz & Lahiri, 2004) by using the Mismatch Negativity (MMN) as event-related-brain potential. The MMN is an automatic change detection response in the brain (Näätänen, Gaillard, & Mäntysalo, 1978; Schröger, 1998; Näätänen 1992; Näätänen, Paavilainen, Rinne, & Alho, 2007 for a review) that is suggested to reflect long-term memory traces for language sounds such as phonemes (Dehaene-Lambertz, 1997; Näätänen, 2001; Näätänen, Lehtokoski, Lennes, Cheour, Huotilainen, et al., 1997), syllables (Shtyrov, Kujala, Palva, Ilmoniemi, & Näätänen, 2000) and lexical representations of words (e.g. Jacobsen, Horváth, Schröger, Lattner, Widmann, & Winkler, 2004; Pulvermüller, Kujala, Shtyrov, Simola, Tiitinen, et al., 2001; see review Shtyrov & Pulvermüller, 2007b; Pulvermüller & Shtyrov, 2006a) and even to fine-grained featural differences of speech sounds such as place-of-articulation or voicing differences (e.g. Kraus, McGee, Sharma, Carrell, & Nicol, 1992, Kraus, McGee, Carrell, Sharma, Micco, & Nicol, 1993a; Aulanko, Hair, Lounasmaa, Näätänen, & Sams, 1993; Phillips, Pellathy, Marantz, Yellin, Wexler et al., 2000a, Phillips, Pellathy, & Marantz, 2000b; Phillips, 2001; Maiste, Wiens, Hunt, Scherg & Picton, 1995). MMNs are obtained by presenting the subject with occasional infrequent, ‘deviant’ stimuli after a series of identical, ‘standard’ stimuli. It appears as a negative deflection from the onset of the deviation around 100-250 ms at fronto-central electrode sites.

The use of the MMN is an excellent measure to study mental representations and assumed differences in speech sound representation, since it is sensitive to language sounds and featural differences of sounds. It is assumed that different processes are active between the standard and the deviant. In order for our perceptual system to recognize that a deviant is different from the standard there must be a memory activated by the standard. The activation of a differential response by the deviant involves a process of comparison with the memory trace of the standard, signaling a mismatch between the stored memory trace and the physical features of a deviant stimulus. Therefore, the MMN is elicited when the sensory

input does not match with the memory trace activated by the preceding standard stimuli (Näätänen, 1992).

Following this logic, **Eulitz and Lahiri (2004)** proposed that the series of identical (standard) stimuli form a central sound representation where the information structure is close to the format used for long term memory (e.g. Cowan, Winkler, Teder, & Näätänen, 1993; Näätänen, Schröger, Karakas, Tervaniemi, Paavilainen, 1993), i.e. close to the so called underlying representation in the mental lexicon. An infrequently presented (deviant) stimulus creates a sound percept, which corresponds to the surface representation formed by phonological features that are extracted from the acoustic signal. When standard and deviant come together in the MMN design the surface representation of the deviant is compared to the underlying representation of the standard.

Thus, the observed change detection response (MMN) reflects – besides the acoustic change – the comparison between the surface and the underlying representation. This makes the MMN a very good measure to investigate how the acoustic signal is matched to representational information, which seems vital for representational models such as FUL. Eulitz and Lahiri (2004) investigated three vowels in their MMN study, the coronal vowels [e], and [ø], as well as the dorsal vowel [o]. When [e] and [ø] are presented as deviants it is assumed that they extract [CORONAL] from the acoustic signal. They are compared against a row of /o/ stimuli as standards which have formed a central sound representation pre-activating [DORSAL] from the lexical representation. Following the FUL model this creates a conflict, since [CORONAL] is extracted from the deviant and maps onto the stored representation of the feature [DORSAL]. In contrast, if the standards and deviants are reversed, i.e. /ø/ and /e/ as standards preceding the deviant [o], then no conflict is expected. [DORSAL] is extracted from the acoustic signal of the deviant [o] and mapped onto an empty space of the underspecified [CORONAL] activated by the standards /e/ or /ø/. Eulitz and Lahiri's (2004) crucial result were the following (Figure 3): The same acoustic standard - deviant contrasts triggered asymmetric MMNs when they were reversed in their functions of standard and deviant in the coronal ~ dorsal contrast. Earlier latency and higher amplitude MMN values were found for the coronal vowel deviants [e] and [ø], when [DORSAL] was pre-activated by the standard /o/, which seemed to have created a conflict. Smaller amplitudes were found for the reversed non-conflicting case, when the coronal vowels served as standards. However, the coronal vowel contrast [e] and [ø] did not differ in amplitude or latency independent of whether they served as a standard or deviant (Figure 3). These differential MMN asymmetries for similar acoustic/phonetic differences between pairs of vowels in isolation were discussed as a reflection of the brain referring to underspecified phonological representations.

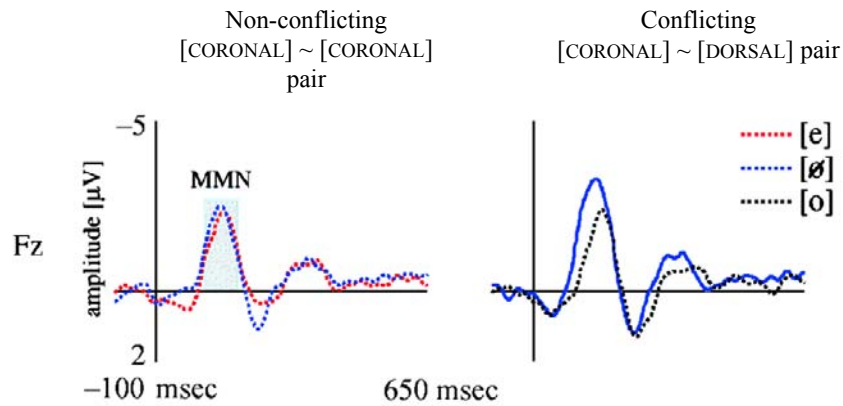


Figure 3: MMN waveforms for all vowel pairs reversed as standard and deviant are shown. The color of the waveforms indicates the deviant vowel, red for [e], blue for [ø], and black for [o]. Solid lines indicate the only conflicting condition and dotted lines, the non-conflicting conditions (Adapted from Eulitz & Lahiri, 2004).

Further neurophysiological evidence for underspecified representations comes from another Mismatch Negativity study by **Walter and Haquard (2004)** using the magnetencephalogram (MEG). Their subjects were presented with syllable initial nasal and plosive consonant-vowel segments (CVs) each serving as a standard and deviant. In the plosive condition the feature [CORONAL] of the syllable [da] was contrasted with the [LABIAL] of [ba], and in the nasal condition [CORONAL] of [na] was contrasted with the [LABIAL] in [ma]. Following FUL and the logic used in the Eulitz and Lahiri (2004) study a stronger mismatch field response (MMF)¹¹ would be expected when a coronal deviant is compared to a labial standard stimulus, where [CORONAL] is extracted from the signal mapping onto the pre-activated stored representation of the feature [LABIAL] which is assumed to conflict. In the reversed case, a labial deviant is compared against a coronal standard, thus a weaker MMF response is expected since no conflict is created due to the underspecification of [CORONAL]. The results confirmed this prediction: In a late MMF time window of 200-300 ms the coronal deviants elicited a higher MMF amplitude (& earlier latency as tendency) compared to the labial deviants. These acoustically equidistant pairs evoked different mismatch responses depending on the order in which they were presented. The authors concluded that these differences were due to comparisons being made between representations differing not only in the specification of place of articulation, but in whether such specification is present at all. Degrees of phonological specification seem to exist which are relevant in speech processing.

¹¹ Magnetic mismatch field (MMF) equivalent to the MMN.

In summary, many studies have shown that underspecified representations such as the place feature [CORONAL] seem to be a fundamental principle in the organization of the mental lexicon as they can deal with the variation and asymmetrical surface realizations that occur in natural speech. The studies reported here have focused on place of articulation finding evidence for the underspecification of the feature [CORONAL] for consonants in final and medial position in isolated words and sentential contexts (Wheeldon & Waksler, 2004, Lahiri & Reetz, 2002; Friedrich et al., 2006; Friedrich et al., 2008), in syllable initial position (Walter & Haquard, 2004), as well as for isolated vowels (Eulitz & Lahiri, 2004). One study examined manner of articulation in vowels cross-linguistically (Lahiri & Marslen-Wilson, 1991). Further experiments have been conducted investigating height features (Felder, 2006; Felder et al., 2008; Scharinger et al., 2009; Scharinger & Lahiri, 2010) and features that have been deleted or reduced in production and comprehension studies (Zimmerer et al., 2009; Zimmerer, 2009).

This dissertation will further test place of articulation in linguistically more complex structures such as words and nonwords, as well as manipulating place in different sounds such as plosives and nasals. The focus of this work lies on new feature dimensions such as manner of articulation, contrasting features such as [NASAL] and [STRIDENT] or [PLOSIVE] with different speech sounds, as this still is untested territory.

1.3.3 *Contradicting Evidence*

The experimental evidence however is not consistent in supporting underspecification in lexical entries. **Gaskell and Marslen-Wilson (1996)** for instance found mixed support for underspecified representations in his cross-modal repetition priming experiment. Sentences were presented auditorily containing a prime word in either phonological changed or unchanged form. At the offset of the prime, a visual target was presented being always the intact test prime (unchanged form) and subjects had to make a lexical decision response. The prime words ended either in their original unchanged form with a coronal segment (e.g. '*We have a house full of fussy eaters. Sandra will only eat lean bacon*'), or ended with a word-final change to a non-coronal segment (e.g. '*...will only eat leam bacon*'). In a control condition a completely different prime word was used (e.g. '*...will only eat brown loaves*') in order to provide a baseline against which priming effects can be evaluated. In a first experiment the test sentences were presented up to the offset of the prime word, making the following context of the primes unavailable (e.g. *leam* was followed by the visual target *LEAN*). Results showed equivalently strong priming effects in both the changed and unchanged conditions, which indicate that even the distorted words activated their lexical forms. In a second experiment the prime sentences were presented embedded in full

sentential context, allowing the following context of the phonological changes to affect the matching process. The word following the prime either licensed the change of the final consonant in terms of assimilation (viable context: ‘...will only eat *lean* bacon’) or not (unviable context: ‘...will only eat *lean* gammon’). Results showed faster reaction times for changed and unchanged conditions in the viable context, but no priming for the changed condition in an unviable context. Thus, changed primes presented in a phonologically viable context for assimilation produced no mismatch compared to the changed primes in the unviable context where priming was reduced considerably. The inappropriate context (*lean* gammon) did not license the final [LABIAL] in *lean* as variant caused by assimilation to activate the underlying form *lean*.

Results of the first experiment showed equivalent strong priming effects in both, the changed and unchanged conditions, which indicate that even the distorted words, activated their lexical forms. This speaks for underspecified representations as proposed by the FUL model, where the noncoronal pseudowords activated their original base words. Here the context did not play a role. However, in the second experiment, the following context of the phonological changes did affect the matching process. Results showed a significant difference between changed and unchanged primes in inappropriate context, but not in appropriate context. These results contradict the findings by Wheeldon and Waksler (2004) who did not find any context effects. An underspecification theory as provided by FUL would not account for context having any influence on priming whereas phonological inference rules as proposed by Gaskell and colleagues would.

A more recent study by **Gaskell and Snoeren (2008)** examined the degree of assimilation as well as how context affects lexical access. They used sentences in which the target consonants may or may not lead to assimilations depending on the following context (e.g. ‘*I think a quick run/rum picks you up*’ and ‘*I think a quick run/rum does you good*’). In their first perception test investigating the degree of assimilation, subjects had to choose between pairs of target words such as *run-rum*. They perceived significantly more non-coronals (*rum*) when the underlying coronal was followed by a non-coronal (e.g. *picks*), allowing assimilation to happen but not vice versa. The presence of assimilatory context caused ambiguity.

In a second experiment they presented only non-coronal targets (e.g. *rum*) in a neutral and semantically biasing preceding context, and additionally cross-splicing the following context. The targets were presented in a viable context (e.g. *rum picks*) with a following non-coronal, or a non-viable context (e.g. *rum does*) with a following coronal. Subjects were required to identify the target word in a two-alternative selection task. Presented with the non-coronal words subjects heard more coronal targets (*run*) when they

were presented in a viable context than in a non-viable context. This implies that they activated the underlying coronal in contexts where assimilation would license the non-coronal. When a preceding sentence biased the percept semantically towards coronals as in *run* (e.g. ‘*It’s best to start the day with a burst of activity. I think a quick rum picks you up/does you good*’) the coronal responses increased but still showed the same asymmetry: more coronals were perceived in a viable context (following labial) than in a non-viable context (following coronal). This asymmetric pattern changed to the reversed when the following context was cross-spliced. Now, there were more coronal responses to a viable context (e.g. *rum_{does}picks*¹²) but where the target originally preceded a non-viable context. Less coronal responses were perceived in the non-viable context (e.g. *rum_{picks}does*¹³) where *rum* originally preceded *picks*. This cross-splicing showed that the same acoustic stimulus triggered different responses depending on the following context.

Their first experiment is consistent with an underspecification account, since listeners hardly gave labial responses (*rum*) upon hearing an underlying coronal (*run*) independent of the context. According to FUL, when extracting [CORONAL] from the signal in the case of *run* it mismatches with [LABIAL] rejecting words such as *rum*. The results of the second experiment however crucially depended on the context. A noncoronal surface-form such as *rum* actually was perceived as *run* more often when the following context was viable (*rum picks*) compared to a nonviable (*rum does*). Gaskell & Snoeren (2008) argue that these results seem more challenging for an underspecification account since contextual influences should be irrelevant to the matching process. Lahiri and Reetz (2010) nevertheless argue that these results do not necessarily contradict with underspecification since FUL’s basic assumptions are not violated: coronal and non-coronal responses were about 50:50 when hearing an underlying labial. When a [LABIAL] feature is extracted it can activate underlying labials and coronals. Therefore, hearing *rum* activates *rum* and *run* to about 50 %. The context effects slightly change the results facilitating the response choice and making recognition easier for the perceptual system, however recognition is not disrupted by changing the context.

Furthermore, **Gow (2002)** conducted a series of cross-modal form priming experiments where word final consonants of prime words changed into other existing words (e.g. *right* - *ripe*) creating potential lexical ambiguity, in order to test the pattern of lexical activation after hearing ambiguous words in contexts that support assimilation. In each experiment, subjects either heard an assimilated prime word embedded in a sentence context

¹² *rum_{does}picks*: *rum* from the original *rum does* phrase was cross-spliced to *picks*, resulting in *rum picks*.

¹³ *rum_{picks}does*: *rum* from the original *rum picks* phrase was cross-spliced to *does*, resulting in *rum does*.

that licensed it by inducing labialization of the final segment (e.g. *This time she tried to get the right herries for her pie*) or they heard an unassimilated prime word which might sound like an assimilated form (e.g. ‘*This time she tried to get the ripe herries for her pie*’). After hearing the prime a lexical decision target word was shown, in the original coronal form (e.g. *RIGHT*), a possibly assimilated noncoronal form (e.g. *RIPE*), or a phonologically and semantically unrelated item (e.g. *HEM* or *HEN*).

The results revealed that the underlying coronal primes activated coronal targets, but not noncoronal targets. This means that when subjects heard an assimilated token of the word *right* that perceptually approximated both *right* and *ripe*, only *right* was accessed, but not *ripe*. This finding is not consistent with predictions made by FUL, which assumes that listeners would simultaneously access both *right* and *ripe*, since the underspecified representations of these items would not mismatch with an assimilated item of *right*.

Gow interprets the results as indicating that assimilation produces an incomplete or non-categorical change in the acoustic cues to place features, which leaves the listener with enough information (coronal) to determine the non-assimilated value of the place feature even in strongly assimilated items, thus accessing only the original form *right* rather than *ripe*. He argues that feature cues are associated with neighboring segments showing overlapping temporal distributions. Assimilated items show a pattern of formant movement that can be intermediate between the patterns associated with unmodified coronal and noncoronal place. Assimilation can produce a range of degrees of acoustic information. Gow argues that the acoustic signal provides sufficient information to disambiguate the underlying form of a word. Place assimilation seems to provide enough acoustic information to directly access its unmodified representation. This however, does not contradict with FUL. When indeed [CORONAL] is left in the signal it will activate *right*, but will mismatch with the noncoronal *ripe*, which is what Gow found.

In Gow’s third experiment, the identical carrier sentences were used, but now the word *right* was replaced by the original word *ripe*. Here the prime items ended in acoustically unambiguous noncoronals. It was found that *ripe* only primed *ripe*, even showing considerable negative priming for *right*. Gow argues that these results are inconsistent with underspecification.

In his fourth experiment the context that drove the assimilation was removed and showed a pattern that is consistent with the predictions made by FUL. Strongly assimilated primes of the word *right* were only used in a sentence fragment (e.g. ‘*This time she tried to get the right /ʔ*’). This time the assimilated form of *right* primed both the coronal form *right* and the noncoronal form *ripe*. It seems that context plays a role in disambiguating the correct forms for lexical entry. It seems that Gow’s experiments do not show a clear picture for or against FUL.

Further contradictory evidence comes from **Mitterer and Blomert (2003)**, who investigated whether perceptual compensation for phonological assimilation can be found in viable or unviable changed phrases. In order to test the perception of assimilated tokens, Dutch words were presented in their canonical form such as *tuin* ‘garden’ or with a place change for the final nasal (**tuim*) in isolation or in a phrase that did (e.g. *tuinbank* ‘garden bench’; viable condition) or did not allow assimilation (e.g. *tuinstoel* ‘garden chair’; unviable condition). A two-alternative forced choice task was used, in which subjects had to decide whether the word *tuin* was pronounced canonically or with the word-final [m] as in **tuim*. Identification performance was near ceiling in both the no-context and unviable context conditions. In the viable condition however, more errors were made, showing a significant trend to (mis)perceive **tuimbank* as *tuinbank*. These results so far do not contradict with FUL.

To rule out attentive and decision based processes they further ran an MMN study. They tested the perceptual distance between */tuinbank/* as standard and [**tuimbank*] as deviant in a condition where the change from the [n] to [m] is viable, and compared this to a condition, where the change is not viable (e.g. standard */tuinstoel/* and deviant [**tuimstoel*]). A significantly larger MMN was found in the unviable context compared to the viable context. The viable context did not elicit an MMN response at all. These results indicate that a canonical phonological form (e.g. *tuin* in *tuinbank*) and a viable change (e.g. **tuimbank*) are perceptually less distant than a canonical form and an unviable change (e.g. **tuimstoel*). The authors interpret this as strong evidence for the existence of regressive context effects in the perception of place assimilation at a prelexical level.

FUL would not predict a difference based on context as it was the case here. However, a representational account like FUL could only be tested if all these conditions were also tested in the reversal of standard-deviant contrasts as in Eulitz and Lahiri (2004) in order to tap into representational differences, which was not done in this experiment.

This was tested in another MMN experiment by **Mitterer (2003)** who used the same Dutch words as above, *tuin* with a word final coronal and its possible assimilated pseudoword with a final labial as in **tuim*, both serving as standard and deviant in an MMN design. According to FUL the deviant [tuin] would lead to a mismatch with the pseudoword standard **tuim/* since it conflicts with its underlying [LABIAL] representation of /m/ (conflict condition), while the deviant [**tuim*] would not mismatch with the standard */tuin/* due to the underspecified [CORONAL] of /n/ (non-conflict condition). Thus it was expected to find a difference in MMN pattern when the roles of standard and deviant change. However this is not what Mitterer found: both MMNs did not differ. These results speak against underspecified representations favoring the assumption of fully specified lexical

representations and pre-lexical compensation mechanisms in dealing with variation. Nevertheless, it should be noted that there was a trend for a latency difference between the conditions, showing an earlier MMN onset for the condition with the phonological conflict similar to the results of Eulitz and Lahiri (2004), which is not mentioned in the study. Also the results are based on only a very few subjects ($n = 9$) which might have led to these weak effects.

It is to note that the compound stimuli of these studies by Mitterer rely on context effects where transitional probabilities seem to play a role. If *tuin* and *bank* co-occur frequently, the underlying coronal nasal may require labiality by virtue of this co-occurrence. The question though concerns the level of representation. FUL does not assume transitional probabilities on the level of representations. But nothing would speak against transitional probabilities in compounds or high frequency phrases on higher cognitive levels as is proposed for instance by Shortlist (Norris & McQueen, 2008). On a higher level, top-down expectations may be generated, and reflect assimilation processes. This is fundamentally different from claiming that the underlying nasal actually is labial. The probability of having certain sounds in certain contexts may affect perception; however, this does not change the underlying representations. When asking subjects about the last consonant in **tuimbank*, one will most certainly elicit the answer [n] and not [m] (also see perception study by Mitterer & Blomert, 2003). This knowledge reflects the long-term memory of these sounds, in total accordance with FUL. Therefore the effects described by Mitterer may play a role at a different level other than the representational level.

These results provide ambiguous answers in what influences lexical access. Many factors obviously seem to influence the process of lexical activation. For instance semantic, syntactic, and prosodic information, also experimental factors such as prime type (e.g. word or nonword, semantic, identity, or fragment) and task related aspects such as decision and attentional biases, segment change and phonetic detail in the acoustic signal, sentential context or isolated words and segments can all influence the outcome of these studies.

1.3.4 Summary

The human ability to recognize speech is a remarkable robust characteristic of spoken language processing despite the high variability and disruptive factors that occur while the speech stream is unfolding over time. Our brain and perceptual system has to deal with mispronunciations, speech errors, masking of speech sounds due to environmental noise and acoustic-phonetic variability such as assimilation. For example, the adjective *blonde* can be pronounced in different ways depending on the following context, for instance [blɔŋ] in

blonde girl, [blon^d] in *blonde Daisy*, or [blom] in *blonde boy*. These variances however hardly seem to interfere with ordinary perceptual interpretation and lexical access.

In recent years, increasing interest has grown as to how the incoming acoustic signal is mapped onto mental representations in the mental lexicon to activate words and meanings. One can distinguish at least four different approaches to the problem of phonological variation and how to cope with the fact that the same word can be realized with different surface forms. The first three approaches all take the view that the core representation of a lexical form is a fully specified representation. In the first approach, lexical entries are made up of very detailed memories based on the experience, which are continuously updated and comprise all surface variants of a given word (Bybee, 2001; Johnson, 1997). Linguistic categories are modeled as collections of experienced instances, rather than seeing speech perception as a process of mapping variable input onto invariant abstract representations.

A second approach assumes a single canonical representation of each word that is activated by a lexical search mechanism. Phonemes or phoneme probabilities are provided as input that start a lexical search activating the best matching representation (Norris, 1994; Norris and McQueen, 2008). Probability computations and likelihood functions can account for variable and noisy speech input, as well as asymmetries that arise due to assimilation processes.

A third approach, also supporting the assumption that the target representation is fully specified proposes an active process of compensation that mediates between the input string and the lexical level (Gaskell et al., 1995; Gaskell & Marslen-Wilson, 1996, 1998; Gaskell and Snoeren, 2008). Phonological inference rules allow regularly viable surface forms to be mapped onto fully specified representations. For instance, Gaskell et al., (1996) found that only assimilated variants changing into a pseudoword (e.g. **leam*) of a base word (*lean*) accessed the lexicon if the assimilation occurred in a viable context (**leam bacon*), but not in a non-viable context (**leam gammon*). A more recent study by Gaskell and Snoeren (2008) found that the same acoustic word stimuli trigger different responses depending on the context.

The fourth approach differs from the other three by making entirely different assumptions about the underlying mental representation of the lexical form, and solving the problem of phonological variation by making explicit assumptions about the abstractness of lexical representations (FUL; Lahiri & Reetz, 2002; 2010). These representations are made up of abstract hierarchically structured features operating on the basis of the mapping process between the auditory input and a stored representation in memory. Importantly only distinctive, non-redundant information is represented and specified such as the features [LABIAL] and [DORSAL]. Features such as [CORONAL] being considered the default place of

articulation are underspecified and therefore are likely to undergo change. For instance, the final plosive in *blonde* is underspecified for its place of articulation [CORONAL]. Consequently, [blom], [blon^d] and [blon] would all be possible variants of the underlying partially specified /d/ in the mental representation. Segments with [LABIAL] or [DORSAL] rarely assimilate to other places of articulation, since they are assumed to be fully specified.

This account of underspecification of certain features explains asymmetries that occur naturally in phonological variation. The feature representation in FUL is not just based on local contexts, but on a universal system of contrasts. One major difference between the other models and FUL in dealing with phonological variation is that the other approaches rely on experience (frequency of occurrence) or contextual information.

Meanwhile, there is a considerable amount of literature finding behavioral and neurophysiological evidence for the underspecification of [CORONAL] in perception and production (e.g. Lahiri & Reetz, 2002; Friedrich et al., 2006; Friedrich et al., 2008, Eulitz & Lahiri, 2004; Zimmerer et al., 2009; Walter & Haquard, 2004, Obleser, et al., 2003; Gumnior, et al., 2005; Wheeldon & Waksler, 2004). For example, Wheeldon and Waksler (2004) found processing asymmetries between [CORONAL] and [DORSAL] place of articulation, arguing for the underspecification of [CORONAL] in the underlying representation without an effect of context. Furthermore, it has been shown that the underspecification of [CORONAL] is not specific to a certain word position. Asymmetric activation patterns between [CORONAL] and [DORSAL] have been found word medially in ERP studies (Friedrich et al., 2006) and in initial position (Friedrich et al., 2008) where this degree of abstractness in the representation plays a role in word recognition. Additional evidence was found for isolated, context-free vowels (Eulitz & Lahiri, 2004) and consonants (Friedrich et al., 2006; Friedrich et al., 2008; Walter & Haquard, 2004). Eulitz and Lahiri (2004) found an enhancement of the mismatch negativity, as well as an earlier latency when the vocalic place representation was specified, rather than unspecified in the case of [CORONAL] (e.g. /o/ [DORSAL] ~ /ø/ [CORONAL]).

1.3.5 Research Question

The main question remains. How much detail must there be in the acoustic signal and in the lexical representation for the lexicon to recognize words and variants of words? How is the lexicon accessed and the appropriate sound or word selected?

So far, different solutions have been proposed by the different models described above. In this work the focus lays on sparse and abstract underspecified feature representations in dealing with lexical access as proposed by the FUL model. FUL seems to be able to account for many problems and phenomena respectively, along the mapping from

the acoustic speech signal to its mental representation. It is context-independent, relies on universal and language-specific principles, and can explain asymmetries due to assimilation processes and other variances in the speech signal quite easily. To further test the assumptions of FUL, several experiments have been conducted in this line of research, where syllables, words and nonwords were used in a mismatch negativity design similarly to the study by Eulitz and Lahiri (2004).

This dissertation is aimed at generalizing the previous results of underspecified mental representations for place-of-articulation features to further speech sounds and other featural dimensions such as manner of articulation. It is of interest whether the brain indeed refers to underspecified phonological representations as a fundamental principle of the mental lexicon in dealing with language variation and mapping the speech signal onto mental representations for lexical access. A representational model such as FUL seems quite accurate with respect to the mapping from acoustic to representational information. This mapping appears to be a key process in the perception of spoken language. A good method to investigate this matching process as well as the representational level is the MMN as brain response.

This thesis employs event-related brain potentials as the main technique because it is an excellent tool for studying language processes online. In particular, the MMN will be used throughout all main experiments as it captures early and fast automatic sensory processes which are relevant for this work. How the mapping from the acoustic detail onto phonological representations takes place and how detailed these representations are stored, are of main interest. The next chapter (1.4 ‘Methodology’) will describe this electrophysiological measure and the MMN in greater detail.

The first MMN experiment in this dissertation tries to replicate and extend the findings of Eulitz and Lahiri (2004) for isolated vowels, by using the same three vowels contrasting the place features [CORONAL], and [DORSAL], but embedding them into linguistically more complex structures such as words and nonwords. The second MMN study attempts to expand the findings of specified and underspecified place-of-articulation features to other speech sounds, using three syllable initial stop consonants and contrasting the features [LABIAL], [CORONAL], and [DORSAL]. The following MMN experiments try to generalize the underspecification account to other feature dimensions such as manner of articulation, using meaningless VCV-syllables. Experiment three makes a first step by manipulating place features in a [CORONAL] ~ [DORSAL] contrast, as well as manner features in a [NASAL] ~ [STRIDENT] contrast. It was of interest whether the MMN is sensitive to fine-grained manner feature differences at all. Experiment four further generalizes the specification of representations by exploring the manner feature [PLOSIVE], testing its assumed underspecification and contrasting it with the specified feature [NASAL]. The fifth

experiment aims at finalizing and extending the assumptions made for underspecified and specified manner and place features with additional speech sounds such as nasals, plosives, coronals and labials. The sixth experiment is a behavioral discrimination experiment to control for acoustic and attentional effects of all the consonantal sound contrasts used in the MMN studies.

1.4 Methodology: Possibilities for Studying the Mapping from Sound to Representation

1.4.1 Electroencephalography – EEG

For over a century, scientists have been using measures of brain activity to gain insights into perceptual, cognitive, and motor functions. Researchers have developed a variety of methods to noninvasively measure brain activity, one of which is an electromagnetic approach, examining electrical and magnetic activity in both temporal and spatial domains (e.g. Näätänen, Ilmoniemi, & Alho, 1994). One of these approaches, the electroencephalography (EEG) provides a continuous recording of overall brain activity, and has proven to have many important clinical and scientific applications. Electrical brain activity was first recorded in rabbits by Caton in 1875; the human EEG was first described by Berger in 1929. Berger showed that one could measure the electrical activity of the human brain by placing an electrode on the scalp, amplifying the signal, and plotting the changes in voltage over time (Berger, 1929; cited in Luck, 2005). It is a non-invasive technique with a high temporal, but low spatial resolution for evaluating brain activity underlying different brain functions.

1.4.1.1 Neural and Electric Basis of the EEG

The neural activity, measured at the surface of the scalp is an electrochemical process, arising from neurons in the cerebral cortex and is assumed to reflect the extracellular current flow. This includes activity that arises spontaneously or in response to sensory stimuli. The current flow recorded by the EEG is associated with the summation of *post-synaptic potentials* in synchronously activated, vertically oriented pyramidal cells in the cortical layers¹⁴.

The proportion of currents that flow through the extracellular space is responsible for the generation of electric fields, which can be recorded by the EEG. To measure an electric field on the scalp a large number of neurons in the order of 10^3 - 10^4 cells have to be active in synchrony and need to be aligned in parallel to cause a detectable *open electric field*. This organization leads to the summation of the associated currents. If neurons are not arranged in parallel the result is a *closed field*. It has the distribution of a sphere with zero potential and no field is generated outside. Therefore, the electric activity of such closed fields cannot be

¹⁴ The condition of parallel alignment and synchronized activation determines pyramidal cells as the origin of the EEG. They are organized parallel in the outer layers of the neocortex reaching into deeper layers and connecting to other cortical areas with their dendrites. They represent 75 % of neurons in the cerebral cortex (Braitenberg, 1977). It is suggested that the scalp recorded waveforms are a reflection of post-synaptic dendritic potentials, rather than of axonal action potentials (Allison, Wood, & McCarthy, 1986; cited in Rugg & Coles, 1996)

recorded at the scalp. Positivity or negativity at the scalp reflects the direction of current flow between the active neurons and the recording site.

The surface electrodes measure the change in voltage corresponding to the difference in potential between the signal at a recording electrode and that at a reference electrode. In general the measured brain activity is described with regards to different frequencies (0.5 and 50 Hz) and amplitudes (1 and 200 μV) as parameters, which depend on developmental conditions as well as on activation states. A classification of mainly four frequency ranges has been established in order to describe the spontaneous EEG (Cooper, Osselton & Shaw, 1984).

The plain EEG however is limited in providing insights to cognitive processes because the recording tends to reflect the brain's global electrical activity. A more powerful approach used by many neuroscientists focuses on how brain activity is modulated in response to a specific task or sensory stimuli, known as 'evoked potential' or 'event-related brain potential'.

1.4.1.2 Measuring EEG and Event-Related Brain Potentials (ERP)

To measure the EEG, pairs of electrodes are attached to the surface of the human scalp with a conducting paste and are connected to a differential amplifier; the output of the amplifier reveals a pattern of variation in voltage over time, which is known as the EEG. The amplitude of this voltage variation varies between about -100 and +100 μV , and ranges in its frequency up to 40 Hz or more. The EEG is measured as relative difference of potentials across time between two electrodes. The electric potential of one electrode (or of each electrode of an array of electrodes) is recorded compared to a 'reference' electrode set at an electrically relative inactive position (e.g. mastoids, nose, ear lobes 'monopolar recording'), or compared against an active electrode over a different scalp area (comparing differences between subsets of two active adjacent electrodes 'bipolar recording') or even more complex by subtracting the across electrode mean voltage from each electrode to yield recordings with respect to an 'average reference'.

Small voltage changes in the EEG, a tiny signal embedded in the ongoing electrical activity of the brain which follows or precedes some internal or external event with a relatively constant time course are called event-related potentials (ERP). They occur time-locked or synchronized in response to a sensory, cognitive, and motor event and provide direct and precise information about the time course of information processing.

The most common method to extract ERPs from the EEG is the signal-averaging technique of many EEG samples that are time-locked to the repeated occurrence of a particular event. Given the assumption that EEG activity that is not time-locked to the event

will vary randomly across periods (50-100 μV) fluctuating between 8 to 12 Hz in the silent wakeful state (Berger, 1929) this ‘background’ noise will be reduced by averaging, leaving the event-related activity. This averaged waveform is very small (1-30 μV) relative to the background and should largely represent activity that carries a fixed temporal relationship to a sensory, motor, or cognitive event across epochs. After extracting the signal, the voltage by time function results in a waveform with a number of positive and negative peaks. These peaks are described as components reflecting different sensory, cognitive (e.g. evaluation processes) and motor processes that are classified on the basis of their scalp distribution and response to experimental variables. Luck (2005) defines ERP component as ‘*scalp-recorded neural activity that is generated in a given neuroanatomical module when a specific computational operation is performed*’ (p. 59).

The excellent temporal resolution in the milliseconds range is the most important advantage of the ERP method since it draws a link between brain activity and ongoing behavior. Furthermore, the ERP components are useful measures of covert information processing, as differences between conditions can be obtained in the absence of conscious behavioral responding.

The successful classification of the component is a function of the signal-to-noise ratio, and is determined by the amplitude of a component relative to the background EEG amplitude, the number of experimental trials comprising an averaged waveform, and the degree of artifact in the original data. An ERP waveform can be quantified across four dimensions, polarity relative to a baseline (P – positive or N – negative), amplitude (peak or mean amplitude relative to baseline in [μV]), latency (after stimulus onset as peak latency or component onset in milliseconds [ms]), and topography (spatial distribution across the scalp). Component labels are usually composed of polarity and latency, i.e. N100/P100 (negativity/positivity at 100 ms). By combining temporal and spatial information available in the ERP waveform, it is possible to determine the temporal characteristics (onset and duration) of stimulus- and condition-specific patterns of brain activity.

ERP components are categorized as exogenous or endogenous components. Although this distinction is under discussion and is not entirely dichotomous, components that are dependant on external physical factors of sensory stimuli (i.e. loudness, frequency, luminance, etc.) are defined as *exogenous*, that is, they are not influenced by ‘cognitive’ manipulations and immune to variations in the subject’s state. Components depending more on internal factors and on the nature of the subject’s interaction with the stimulus (i.e. probability of occurrence, task relevance, attention) are defined as *endogenous*. Components – such as the Mismatch Negativity – between 100 and 200 ms after onset of the stimulus

may very well be determined by exogenous, physical and endogenous, psychological characteristics.

ERP research began in 1964, with Walter and colleagues who reported the first cognitive ERP, which they called the *contingent negative variation* or CNV (Walter, Cooper, Aldridge, McCallum, Winter, 1964). The next major advance was the discovery of the P300 component by Sutton, Baren, Zubin, and John in 1965. Over the years, a great deal of research has focused on identifying a variety of cognitive ERP components as well as developing better methods for recording and analyzing ERPs.

In sum, the measurement of ERPs provides a useful method to investigate psychological processes. ERPs can serve as ... 'windows on cognition' – and ... 'windows of the brain' (Coles, 1989, p. 251). Compared to behavioral measures, the advantage of ERPs is that they may directly reflect specific brain functions, whereas most behavioral measures show a final outcome of a complex set of processes and can therefore only be used as indirect indices of psychological processes.

For a more detailed overview about the EEG-method and ERPs it will be referred to introductory literature (Luck, 2005; Rugg & Coles, 1995; Barlow, 1993; Cooper et al., 1984).

In studying language several components can be observed over the course of time, such as the N400 that is sensitive to lexical-semantic integration processes. Two components that are sensitive to syntactic processes are the (early) left anterior negativity ((E)LAN) and the P600.

This dissertation focuses on one ERP component which is not language dependant but sensitive to early language processes, reflecting change-detection in early auditory information processing, the *Mismatch Negativity (MMN)*. A further component that often is followed by the MMN is the P300, which is described as a psychophysiological correlate of attention switching and stimulus-updating. The MMN and the P300 further will be described in the following sections as they will be relevant for this work.

1.4.2 ERP components

1.4.2.1 Automatic Change Detection: The MMN

'It may well be that a physiological mismatch process caused by a sensory input deviating from the memory trace ('template') formed by a frequent 'background' stimulus is such an automatic basic process that it takes place irrespective of the intentions of the experimenter and the S, perhaps even unmodified by the later' (Näätänen, Gaillard & Mäntysalo, 1978, p. 324-325).

The earliest ERP activity that indicates that the brain has detected a change in an otherwise invariant background of homogenous events is called the *Mismatch Negativity (MMN)*. This phenomenon was first described by Näätänen and colleagues (Näätänen et al., 1978) about 30 years ago as a negative component appearing in the ongoing EEG recordings to a tone burst that deviated in frequency from a pattern of identical 'standard' frequency tones. This negativity was elicited for frequency deviance compared to the standards (the same negativity was also elicited for deviating intensities compared to standard intensities). This negativity was elicited for both frequency and intensity deviance compared to the standards for attended and unattended conditions. Näätänen and his colleagues suggested that the MMN reflects the *automatic detection of physical deviance*.

This modality-specific MMN component (and its magnetoencephalic (MEG) equivalent MMNm or the magnetic mismatch field 'MMF') is effectively described as a fronto-central negative component of the auditory event-related potential usually peaking 100-250 ms after stimulus onset. It is elicited by any discriminable change or irregularity in a repetitive characteristic of an ongoing auditory stimulation – irrespective of the subject's attention or behavioral task. It usually shows polarity inversions when recorded from locations below the Sylvian fissure, demonstrating more positive deflections for example at the mastoids. This polarity inversion is due to the location and orientation of the neural source of the MMN (for MMN reviews see Näätänen & Winkler, 1999; Näätänen & Alho, 1995; Näätänen, 1990, 1995; Schröger, 1998; Picton, Alain, Otten, Ritter, & Achim, 2000; Picton, 1995; Näätänen, Paavilainen, Rinne, & Alho, 2007).

The MMN is typically studied in an oddball task, where two classes of stimuli are presented: a frequently occurring 'standard', and an infrequently occurring 'deviant' or oddball stimulus (20% or less). These stimuli are presented to subjects who are awake but are paying attention to something other than the auditory stimuli such as watching a silent movie, or reading a book and ignoring the auditory stimuli (*passive oddball paradigm*). Therefore, the MMN can be elicited in the absence of attention. The MMN can best be

estimated by subtracting the response to the frequent (standard) sound from that of the infrequent (deviant) sound. The MMN is illustrated in Figure 4.

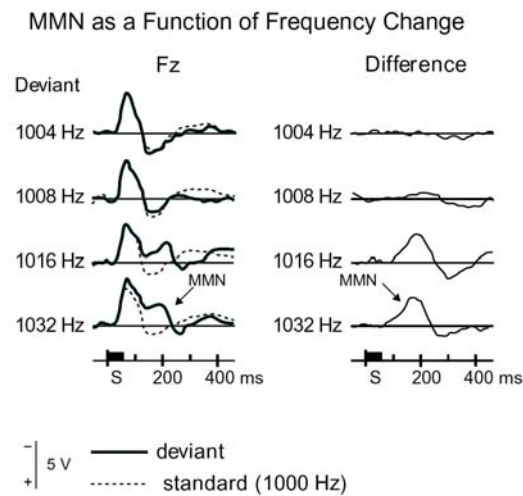


Figure 4: Frontal (Fz) ERPs to randomized standard tones (1000 Hz; 80%; dotted line) and to deviant tones of different frequencies (20%; solid line) as indicated on the left side. The difference waves (MMN) are obtained by subtracting the standard ERP response from that of the deviant stimulus for the different frequency deviants. Note, that the MMN increases with increasing frequency difference between the deviant and the standard. (Adapted from Sams, Paavilainen, Alho, & Näätänen, 1985a; cited in Näätänen et al., 2007)

It has been established that the MMN is generated by a discrimination process, detecting any change in a sequence of sounds. Accordingly, the MMN is elicited when a repetitive sound occasionally changes in frequency (e.g. Sams et al., 1985a; Näätänen et al., 1978), intensity (e.g. Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1987), spatial location (e.g. Schröger & Wolff, 1996), or timbre change (Toiviainen, Tervaniemi, Louhivuori, Lailar, Huutilainen, & Näätänen, 1998). Furthermore, the MMN is elicited by changes in the temporal aspects, such as sound duration, rise time (e.g. Näätänen, Paavilainen, & Reinikainen, 1989; Lyytinen, Blomberg, & Näätänen, 1992), inter-stimulus-interval (ISI; e.g. Sable, Gratton, & Fabiani, 2003), or stimulus order (e.g. Kujala, Karma, Ceponiene, Belitz, Turkkila et al., 2001b). In addition, changes in complex sounds such as chords elicit an MMN (e.g. Winkler & Näätänen, 1993), or even when the deviant changes in more complex stimuli such as phonemes (e.g. Altonen, Niemi, Nyrke, Tuhkanen, 1987; Aulanko et al., 1993; Näätänen et al., 1997; Dehaene-Lambertz, 1997), rhythmic tonal patterns (e.g. Imada, Fukuda, Kawakatsu, Mashiko, Odaka, Hayashi, Aihara, Kotani, 1995) or complex spectrotemporal patterns (e.g. Schröger, Paavilainen, Näätänen, 1994; Näätänen, Schröger, Karakas, Tervaniemi & Paavilainen, 1993; cited in Näätänen & Alho, 1997). The amplitude of the MMN in general increases and its latency is shortened with increasing differences between the deviant and the standard stimuli (see Figure 4). This relationship is

commonly constant, perhaps showing some ceiling effects, as the difference increases (Picton et al., 2000).

These studies demonstrate that the temporal structures of sound patterns are encoded in detail by auditory traces that are involved in the process of change detection. Furthermore, the MMN can be distinguished from other components (e.g. N100, P200, P300), one of which is the N100 waveform that also is associated with the onset of change in an auditory stimulus and appears in the deviant-standard difference waveform. When studying the MMN it is crucial to distinguish both components as they seem to play different roles in perception. The MMN occurs later in time when a particular perceptual response is made, however the N100 can also overwrite the smaller MMN effects. Picton et al., (2000) gives a good overview about the relationship between the N100 and the MMN, and how both components can best be distinguished from one another.

After many years of research, the MMN now has two main paths of research: On the one hand, the MMN is being used with a variety of clinical populations, starting with psychiatric disorders such as schizophrenia (Shelley, Ward, Catts, Michie, Andrews, & McConaghy, 1991; cited in Escera, 2007) and neurological disorders such as Parkinson's disease or comatose patients (Pekkonen, Jousmäki, Kononen, Reinikainen, & Partanen, 1994; Fischer, Morlet, Bouchet, Luauté, Jourdan, & Salord, 1999; cited in Escera, 2007). More recently, it was possible to show specific auditory processing impairments in dyslexia (Baldeweg, Richardson, Watkins, Foale, & Gruzelier, 1999; cited in Escera, 2007), early language development (Kraus & Cheour, 2000) and even for an objective diagnosis of aphasic patients (Aaltonen, Tuomainen, Laine, & Niemi, 1993).

On the other hand, the MMN has been used as a tool for cognitive neuroscience, trying to solve issues on how the brain supports cognitive activities, such as the link between sensory and long-term memory representations (Winkler, Cowan, Csepé, Czigler, & Näätänen, 1996; cited in Escera, 2007), attention (Schröger, 1996), how speech sounds are represented in the left auditory cortex (Näätänen et al., 1997; Dehaene-Lambertz, 1997), including word traces and semantic meanings (Pulvermüller, Shtyrov, Kujala, & Näätänen, 2004; Shtyrov, Hauk, & Pulvermüller, 2004; Menning, Zwitserlood, Schöning, Hihn, Bölte et al., 2005) and even how sounds are organized in auditory memory in processes of auditory scene analysis (Sussman, Ritter, & Vaughan, 1999; cited in Escera, 2007). In this dissertation the MMN is used as a tool for studying mental representations of speech sounds.

1.4.2.1.1 MMN Generation

Since the MMN is not considered to reflect one unitary process, different generators for the MMN are being proposed, (cited in Näätänen, 1995; Näätänen & Michie, 1979; for review see Alho, 1995), mainly two, one in the auditory cortex bilaterally, in particular the supratemporal cortex (superior temporal cortex), the other one in predominantly right-hemispheric frontal areas (for review see also Deouell, 2007).

A major contribution of supratemporal cortical activity to the MMN is elicited by different kinds of stimulus changes as indicated by source localization of scalp-recorded ERPs and magnetic fields (cited in Alho, 1995; Scherg, Vajsar, & Picton, 1989; Giard, Perrin, Pernier, & Bouchet, 1990; Näätänen, et al., 1994). Studies have shown that the MMN reversed its polarity at the mastoids, showing more positive potentials, suggesting bilateral dipolar sources in auditory cortices along the supratemporal plane (anterior part of superior temporal gyrus (STG); Alho, Paavilainen, Reinikainen, Sams, & Näätänen, 1986; cited in Näätänen, 1995). Further evidence for the contribution of auditory-cortex activity to the MMN has been provided by intracranial recordings in animals (King, McGee, Rubel, Nicol, & Kraus, 1995; cited in Näätänen et al., 2007) and humans (Rosburg, Trautner, Dietl, Korzyukov, Schaller, et al., 2005; cited in Näätänen et al., 2007). A number of recordings in humans indicate that the supratemporal MMN sources activated by different kinds of sound changes are anterior to the supratemporal N100 source (Näätänen et al., 2007). These findings suggest that MMNs to different sound changes get a major contribution from areas anterior to the primary auditory cortex (Bertrand, Perrin, & Pernier, 1991). Furthermore, lesion studies have found that patients with anterior temporal-lobe lesions had an attenuated MMN to frequency change, while the N100 amplitude remained within normal limits (cited in Alho, 1995; Mäkelä, Salmelin, Hari, & Hukkanen, personal communication, 1994). Also, Alain, Woods, and Knight (1998) found a selective deficit in processing deviant auditory stimuli presented in the contralateral ear to the lesion in temporal-parietal damaged patients. Imaging studies such as fMRI (Opitz, Mecklinger, von Cramon, & Kruggel, 1999), positron emission tomography (PET; Müller, Jüptner, Jentzen, & Müller, 2002), and event-related optical signals (EROS; Rinne, Gratton, Fabiani, Cowan, Maclin, et al., 1999b) also suggested MMN generation in the auditory cortex (cited in Näätänen et al., 2007). Importantly, there is evidence that simple and complex sounds, such as phonetic changes are processed or represented by different neuronal populations in the auditory cortex (Aaltonen et al., 1993; Näätänen et al., 2007).

In addition to the bilateral supratemporal cortices, other cortical areas are also involved in MMN generation. Patients with lesions in the dorsolateral prefrontal cortex showed reduced MMN by frequency changes in relation to control subjects (Alho, Woods,

Algazi, Knight, & Näätänen, 1994; Alain et al., 1998). An MMN subcomponent generated in the frontal lobe was also suggested by source-localization studies (Giard et al., 1990; Yago, Corral, & Escera, 2001a), from PET (Tervaniemi, Medvedev, Alho, Pakhomov, Roudas et al., 1999; cited in Näätänen & Winkler, 1999) and functional magnetic resonance imaging results (Opitz et al., 1999; cited in Näätänen & Winkler, 1999).

Already Näätänen and Michie, 1979 suggested a functional dissociation between the two generators: the temporal component having a distinctive role in the feature-specific sensory information reflecting either sensory memory storage or a direct comparison process itself. The frontal lobe may be involved in maintaining auditory traces active for comparison with incoming sensory stimuli as orienting response, an involuntary switch of attention to change, either by participating in the memory network, or by facilitating the memory processing top-down. The existence of a coexisting frontal MMN generating network (Giard et al., 1990; cited in Näätänen & Winkler, 1999) fits in with the assumed role of the MMN in initiating involuntary attention switching to the deviating acoustic change occurring outside the focus of attention (Näätänen, 1990).

1.4.2.1.2 MMN & Memory

One important aspect to this research is the relationship between memory and the MMN. The MMN is elicited when the deviant differs from some auditory sensory memory of the previous standards. It only occurs if standard stimuli indeed have previously occurred, which means that a memory trace must have been initiated before a deviance from that memory can be detected. Thus, the memory traces for the repetitive stimulus features that underlie the MMN process are assumed to be part of the long auditory sensory store (Cowan, 1984). However, a repetitive stimulus is not the only possible type of standard that can serve to initiate the MMN process and enter memory representations. It is important that a representation of invariance has been established, a stimulus violating that representation then elicits an MMN. These invariances that build a representation can vary from simple to complex, even being patterned or abstract, where the invariance requires the inference of a relationship ('rule') between the stimuli or their features (e.g. Saarinen, Paavilainen, Schröger, Tervaniemi, & Näätänen, 1992; Picton et al., 2000). For instance, Saarinen et al., (1992) used ascending stimulus pairs as standards at different frequency levels, the second member of the pair always being slightly higher in frequency than the first member. The constant feature of the standards was the direction of the within-pair frequency change. The deviants were tone pairs with frequency change in the reversed direction, namely descending pairs. The MMN obtained was interpreted as a sign of primitive, automatic sensory intelligence, with the standards representing the concept of ascending pairs and the deviants

the concept of descending pairs. Saarinen et al., (1992) suggested that the memory trace appeared to have incorporated a rule or generalization extracted from the stream of physically varying standards.

An MMN therefore can be elicited by violating higher-level regularities of the preceding auditory stimulus stream, indicating that regular inter-sound relationships are represented in the brain and that incoming sounds are checked against these memory representations. The MMN indicates the output of a process that compares incoming activity to a bound and stored representation of the standard stimulus, signaling the detection of this difference (Näätänen & Alho, 1995, 1997; Alho, 1995; Schröger, 1997).

Näätänen (1992) considered the relevant memory to be sensory memory with representations either of the recent acoustic past or of regularities extracted from what is available in sensory memory (Schröger, 1997). Schröger (1997) considered representations of invariance to be stored in the long-term store of auditory memory. Therefore, the MMN generation actually involves two related processes: the maintenance of an auditory memory trace through the standard and activation of a differential response by the deviants through some process of comparison, thus reflecting the output of a comparator that signals a mismatch between the stored memory trace and the physical features of a deviant stimulus (Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995). In other words, the MMN is elicited when the sensory input does not 'match with' the representation or memory trace developed by the preceding acoustic stimulation (of the standards) reflecting a code of stimulus difference or change. This process in particular will be relevant for this line of research.

Näätänen and Winkler (1999) proposed two distinct stages of automatic auditory processing of information, the first being the analysis and extraction of the separate features of a sound. When a sound is presented it elicits an afferent activation pattern in the auditory system, which carries specific stimulus information of sensory features towards higher systems. The later stage is the binding of these features into a perceptible auditory representation, the central sound representation. In this stage, different feature analyzers are integrated so that the developing trace can represent a unitary auditory event. This binding process supposedly uses a sliding temporal window of integration of about 150-200 ms, where temporally closely spaced events are combined into one single perceptual event (Näätänen, 1995; Näätänen, 1990, 2001). The neural populations maintaining the information from the different features of the same stimulus appear to form a single anatomically distributed but functionally integrated and bound stimulus representation.

Since the process generating the MMN is based on these bound stimulus traces formed by previous auditory stimuli the MMN is as an index of auditory stimulus representation (cited in Näätänen & Winkler, 1999; Näätänen & Alho, 1995; Ritter et al.,

1995)¹⁵, that reflects auditory sensory memory. The MMN requires at least two preceding standards before it can be elicited. Cowan and colleagues (Cowan, Winkler, Teder, & Näätänen, 1993) suggested that this sensory memory trace can be in an active or inactive state with regard to MMN elicitation. They observed that a standard-stimulus trace, which had decayed to the extent that the deviant could no longer elicit an MMN, could indeed be re-activated by presenting one standard stimulus. The (dormant) standard-stimulus trace of the previous train was reactivated by the first presentation of the standard. These data suggest that the MMN is obtained only when a previously formed trace is activated.

Further, it is suggested that the sensory information of these traces involved in the MMN generation process might also enter more long-lasting forms of memory (e.g. Cowan et al., 1993; Winkler et al., 1996; Näätänen et al., 1993). Näätänen and colleagues (1993) found a long-term learning effect on the MMN, where subjects were presented with a complex spectrotemporal pattern being repeated at short intervals. The frequency of one segment of this pattern deviated occasionally from that of the standard pattern. For those subjects who were not able to distinguish the deviant at the beginning of the experimental session but learned this discrimination throughout the session did not show an MMN to the deviating pattern at the beginning of the session, but developed a strong MMN in the later part of the session. This rising of the MMN in the course of the session indicates the connection to long-term memory information that forms the basis for the change-detection process. This study together with the study by Cowan et al., (1993) supports the notion that the traces involved in the change-detection process might be linked to long-term representations. This conclusion is also supported by studies showing the MMN reflecting categorical effects in language sound discrimination (e.g. Näätänen et al., 1997; Dehaene-Lambertz, 1997; Aulanko et al., 1993; Phillips, Marantz, McGinnis, Pesetsky, Wexler, & Yellin, 1995; Cheour, Ceponiene, Lehtokoski, Luuk, Allik et al., 1998). Since phonemes are stored in long-term memory, these language studies point to the involvement of the long-term memory information in the traces produced by the MMN. These studies will be discussed in more detail in the next section on MMN and language processes.

To summarize, it has been suggested, that the MMN is generated by a comparison and a neural mismatch between the sensory input of the deviant and the sensory memory representation that has been formed by the standard stimuli.

¹⁵ Näätänen & Winkler (1999) note that the MMN is not a direct expression of auditory stimulus traces. It can rather be used to probe these traces by designing suitable experimental paradigms.

1.4.2.1.3 *MMN and Language Processes*

The MMN as an Index of Memory Traces for Phonemes

Besides the change-detection, and sensory memory processes, the MMN seems sensitive to long-term memory representations activated in the process of language experience. When speech sounds are presented in a passive oddball paradigm, the MMN also is elicited revealing sensitivity to phonetic sounds. The MMN can index auditory discrimination of phonemes (Aaltonen et al., 1987; Titova & Näätänen, 2001; Shestakova, Brattico, Huotilainen, Galunov, Soloviev et al., 2002), indicating that the MMN can reflect the initial stages of acoustic-phonetic processing in the central nervous system. Furthermore, studies have shown that one can access permanent language-specific phoneme memory traces, which are located in auditory cortex, or better the neural generators of the mismatch response can access such phonetic representations (Näätänen et al., 1997; Dehaene-Lambertz, 1997; Phillips et al., 1995). These studies found that native language sounds elicited enhanced MMN responses when they were typical exemplars of a native sound category, than when they were no typical sounds in the subject's native language (e.g. in Näätänen et al's (1997) vowel study: /õ/ of the Estonian language does not exist in the native (Finnish) language). Furthermore, there is evidence suggesting that these language-specific memory traces develop between 6 and 12 months of age (Cheour et al., 1998) demonstrating that learning the phonemes of the mother tongue in infancy brings a corresponding MMN enhancement to phonetically native changes.

Further studies have used cross-linguistic designs to show that the MMN is also capable of indexing phonetic representations in the brain, indexing the detection of phonemes specific to the native language. These language-specific memory representations for phonemes have been found in different cross-language studies such as Finnish-Hungarian (Winkler, Lehtokoski, Alku, Vainio, Czigler, et al., 1999b), English-Japanese (Phillips et al., 1995), English-Hindi (Shafer, Schwartz, & Kurtzberg, 2004; Rivera-Gaxiola, Csibra, Johnson, & Karmiloff-Smith, 2000), French-Hindi (Dehaene-Lambertz, 1997), and French-Japanese (Dehaene-Lambertz, Dupoux & Gout, 2000). These revealed that sound contrasts that cross a phonemic boundary led to larger MMNs than comparable acoustic within contrasts. In addition, Aaltonen, Eerola, Hellström, Uusipaikka, and Lang (1997) found that the MMN corresponds to the perceptual magnet effect (Kuhl, 1991). Learning a second language in adulthood also produced an MMN enhancement (Winkler, Kujala, Tiitinen, Sivonen, Alku, et al., 1999a) showing the formation of new phoneme categories as a result of foreign-language use. These language studies demonstrated that one's native language (and L2) affects the preattentive perception of phonemes. More recent studies have additionally demonstrated MMN sensitivity to phonotactic probabilities (Bonte, Mitterer, Zellagui,

Poelmans, & Blomert, 2005; Bonte, Poelmans, & Blomert, 2007), stress patterns (Weber, Hahne, Friedrich, & Friederici, 2004), audio-visual integration of phonetic information (Colin, Radeau, Soquet, & Deltenre, 2004), voice-onset-time (VOT; Sharma & Dorman, 1999; Phillips et al., 1995), the size of native vowel inventories (Haquard, Walter & Marantz, 2007) and even sensitivity to relatively abstract levels of phonological feature representation (Eulitz & Lahiri, 2004; Phillips, Pellathy et al., 2000a, 2000b; Walter & Haquard, 2004). These results indicate that the MMN appears to reflect the existence of neuronal representations or memory traces for the sounds of languages one is familiar with. When a familiar speech sound is presented, it activates the corresponding phonetic memory trace. This makes the MMN a valuable tool for studying language processes, in particular speech sounds, as is done in this work.

However, there also is some contradicting evidence from studies that have found no MMN effect of phonetic representations (Sharma, Kraus, McGee, Carrell, & Nicol, 1993; Maiste et al., 1995; Sams, Aulanko, Aaltonen, & Näätänen, 1990). Phonetic sound contrasts also always vary in their acoustics. These studies focused on stop consonants and none of them found an increase in MMN amplitude when the acoustic contrast supplemented a phonetic contrast. The amplitude of the MMN within and across sound categories may be due to the varying acoustic differences between the deviant and standard sounds. It was argued that the MMN is not sensitive to phonetic contrasts and can only access acoustic representations of speech sounds. The mechanisms reflected by the MMN therefore may be based on acoustic processing rather than on phonetic processing.

For recent reviews of MMN studies of phoneme perception and representation, see Näätänen et al., 2007, Näätänen, 2001; Dehaene-Lambertz & Gliga, 2004; Pulvermüller & Shtyrov, 2006a; Shtyrov & Pulvermüller, 2007b.

The MMN as an Index of Memory Traces for Higher-Order Linguistic Processes

The MMN is also sensitive to higher-order linguistic phenomena, such as syllables (e.g. Shtyrov et al., 2000; Kraus et al., 1992; Aulanko et al., 1993), morphemes (e.g. Shtyrov & Pulvermüller, 2002a), and words (e.g. Jacobsen et al., 2004; Pulvermüller et al., 2001; Shtyrov & Pulvermüller, 2002b; Shtyrov & Pulvermüller, 2007a; Endrass, Mohr, & Pulvermüller, 2004; Pettigrew, Murdoch, Ponton, Finnigan, Alku, et al., 2004a; Pulvermüller et al., 2004; Shtyrov, Osswald, & Pulvermüller, 2008; Yasin, 2007; for review see Näätänen et al., 2007; Pulvermüller & Shtyrov, 2006a; Shtyrov & Pulvermüller, 2007b).

For instance, Pulvermüller and colleagues (2001) found that the MMN to the same spoken Finnish syllable (e.g. *-ki*) as deviant was larger at the offset of a Finnish word (e.g. *lakki* ‘cap’) than at the offset of a pseudoword (e.g. **vakki*) embedded in a pseudoword context (e.g. **vakke* or **lakke*) in an experiment with Finnish native speakers. These results

indicate that the MMN can reflect the activation of memory representations for native words. Pulvermüller et al. (2001) and Pulvermüller et al., (2004) suggested that the activation of these cortical memory traces for words are realized as distributed and strongly connected neuronal populations linking phonological information mostly stored in the left hemisphere. This lexical enhancement of the MMN was also confirmed by other studies using a variety of languages (Shtyrov & Pulvermüller, 2002b; Endrass et al., 2004; Korpilahti, Krause, Holopainen, & Lang, 2001; Kujala, Alho, Valle, Sivonen, Ilmoniemi et al., 2002; Pettigrew, Murdoch et al., 2004a; Yasin, 2007; for review see Näätänen et al., 2007) For instance, Kujala et al. (2002) found an enhanced MMN for trisyllabic words but not for pseudowords which differed only in one mid-syllable. Endrass et al. (2004) used German function words and pseudowords presenting them in either the left or right ear, or simultaneously in both ears. A greater MMN for words was found in the binaural condition but was absent for pseudowords.

However there is some contradicting evidence as well: Wunderlich and Cone-Wesson (2001) did not find larger MMNs to consonant-vowel-consonant (CVC) words (e.g. *bæd*) compared to consonant-vowel (CV) pseudowords (*bæ*). Also Pettigrew and colleagues (Pettigrew, Murdoch, Kei, Chenery, Sockalingam et al., 2004b) reported the lack of robust MMNs to word stimuli differing only in fine acoustic contrasts, particularly in voiced stop consonants. Shtyrov and Pulvermüller (2002b) argue that the lexical status of the deviant stimulus is relevant for eliciting the MMN, but that the lexical status of the standard stimulus does not affect the MMN amplitude. This contradicts findings that reported ERP responses evoked by word and pseudoword standard stimuli. Diesch, Biermann, and Luce (1998) found that the magnetic MMN (MMNm) for nonword deviants was larger when standards were phonological nonwords than when they were words. Additionally a cross-linguistic study by Jacobsen and colleagues (Jacobsen et al., 2004) found an enhanced MMN for deviants presented in a language-familiar context rather than in a language-unfamiliar context independent of the lexical status of the deviant.

Another interesting finding is that the MMN latency seems to correlate with the stimulus-specific word recognition point suggesting that word-elicited MMNs reflect the processes of individual word recognition by the human brain (Pulvermüller, Shtyrov, Ilmoniemi, Marslen-Wilson, 2006b; Shtyrov et al., 2004). Shtyrov et al. (2004) studied two similar English words (*kick* vs. *pick*) and found similarities between the MMN latencies (30 ms difference) and the word-recognition points (38 ms difference) between the two individual word stimuli. They suggest that the MMN amplitude and latency reveals information about brain processes of lexical access and selection. This body of research showing enhanced MMNs to lexical items indicate that word memory traces activated in the

brain do exist. This presents a unique opportunity to investigate the neural processing of language mechanisms with the MMN.

Additionally, MMN data have also been found to provide evidence for the automatic processing of semantic and syntactic information. It is proposed that MMN sources in perisylvian areas can reflect access to word meaning in the cortex (Shtyrov et al., 2004; Pulvermüller, Shtyrov, & Ilmoniemi, 2005). The processing of semantic features of action words (semantically related to body movements: leg-related word: ‘kick’; hand-related word: ‘pick’) seems to be reflected in the MMN, tapping into semantic processes, similar to lexical access and selection. In a further study, semantic abnormalities causing a semantic mismatch in a sentence (e.g. *Die Frau düngt den *Riesen im Mai*; ‘the woman fertilizes the *giant in May’) was detected by the MMN (Menning et al., 2005) reflecting processes of semantic integration. These results point to the MMN being involved in word recognition processes.

A confirmation of the syntactic role of the MMN comes from the same MMN study violating morpho-syntactic regularities in naturally spoken sentences of German (e.g. *Die Frau düngt den *Rosen im Mai*; ‘the woman fertilizes this *roses in May’; the suffix of the word indexed by an asterisk syntactically mismatches its context). An enhanced MMN amplitude was found for the critical word placed in a grammatically incorrect context (Menning, et al., 2005). Furthermore, a study by Shtyrov and colleagues (Shtyrov, Pulvermüller, Näätänen, & Ilmoniemi, 2003) found an increase of the MMNm for words in ungrammatically context compared to the same word in grammatical correct context. Similar findings were seen in English and German studies (Pulvermüller & Shtyrov, 2003; Pulvermüller & Assadollahi, 2007).

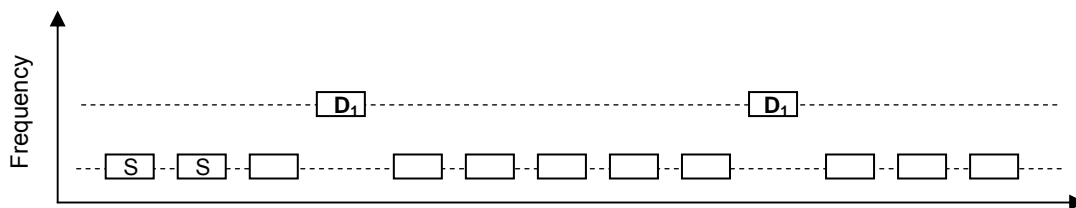
1.4.2.1.4 Variations of the Oddball Paradigm

As mentioned above, the MMN is studied in the oddball paradigm, where a series of auditory stimuli containing high-probability ‘standard’ and low-probability ‘deviant’ stimuli are presented. In recent years different oddball paradigms have been explored. In the *classical paradigm* or constant-standard condition (Cowan et al., 1993), the frequencies of the standards and deviants are fixed throughout the session. Each condition and each type of MMN usually is recorded in a separate block (Figure 5a). Another design, the *roving-standard paradigm* allows for shorter experimental durations but with the same amount of standard and deviant trials as in the classical design. In the roving-standard design, the features of the standards (and deviants) change between stimulus trains. A stimulus train consists of a sequence of standard stimuli, which is followed by a deviant stimulus that is then followed by the next train of standards that are the same as the deviant. The first stimulus of a new stimulus train therefore serves as a deviant for the preceding train, and

forms a new standard category after each deviant (Figure 5b; Cowan et al., 1993; Winkler et al., 1996; Baldeweg, Klugman, Gruzelier, & Hirsch, 2004; Shestakova et al., 2002). Cowan et al., 1993 compared the two designs by testing deviant stimuli in different positions (position 1, 2, 4, 6, or 8) of each stimulus train. No MMN was elicited by deviants at the beginning of a train in either condition, but latest in position 4 an MMN was elicited in both designs. This means that at least three presentations of a standard stimulus seem to create a sufficient basis for the mismatch process between the deviant and the activated memory trace of the standards (Cowan et al., 1993; Winkler et al., 1996) in both paradigms. Figure 5 shows the two different experimental oddball paradigms.

In this thesis both designs will be used depending on the complexity and length of the studies as no differences are expected to surface between the experimental designs. In particular Experiment 2 uses the roving-standard design, whereas Experiment 1, 3, 4, 5 and 6 use the classical oddball design.

(a) Classical Oddball Design



(b) Roving-Standard Design

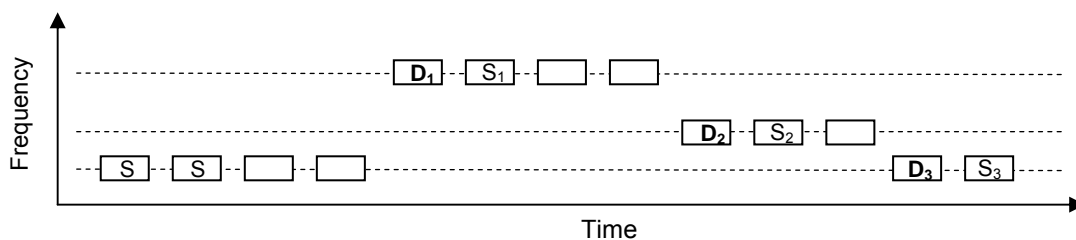


Figure 5: Schematic diagram of the two auditory oddball paradigms. Bars represent the auditory stimuli (e.g. frequencies). (a) Illustrates the classic oddball paradigm with a fixed standard-deviant sequence through out one experimental block. (b) Shows a roving-standard paradigm, where frequencies change through out the experimental session varying different standards and deviants. S = Standard; S_n = Standard with a different frequency; D_1 = Deviant; D_n = Deviant with a different frequency.

1.4.2.1.5 The MMN as Difference Response Controlling the Effects of Physical Stimulus Differences

The MMN reflects a difference waveform that is calculated by subtracting the ERP response to the standard from that of the deviant. In more recent studies, it is not only possible to control for physical stimulus properties in a better way, but also to minimize neurophysiological manifestations of physical stimulus features, including early ERP components such as the N100. Two recording blocks can be run with the standard stimulus in one block becoming the deviant stimulus of the other block, and vice versa. Then subtracting the response to the stimulus when it is the standard from the response to the *same* stimulus when it is a deviant across-blocks results in a same-difference waveform, the ‘identity MMN’ (iMMN). By subtracting the deviant ERP by the identical stimulus presented as standard, thereby eliminating any acoustic differences between the standard and the deviant, it is proposed to obtain the relatively pure contribution of the memory network (e.g. Pulvermüller & Shtyrov, 2006a; Eulitz & Lahiri, 2004; Phillips et al., 2000a, 2000b; Picton, 1995).

The use of the MMN in neurolinguistic studies to understand neural mechanisms underlying speech processes is a powerful tool. Not only is it sensitive to fine-grained perceptual differences, highly automatic, and early in processing, but also reflects the dynamic properties of speech and neural characteristics of the brain. Therefore, it can be applied to study speech sounds which are matched for physical and linguistic features. Using the iMMN approach allows for manifestations of physical stimulus features to be subtracted out at the neurophysiological level and to gain a relatively pure measure of memory representations.

There is one difficulty when studying speech sounds with the MMN, since phonetic contrasts also always differ in acoustic properties. Therefore it often is difficult to determine whether the MMN response can be solely explained on the basis of the acoustic contrast between stimuli or indeed whether phonetic contrasts also elicit a mismatch response. Some studies have tried to circumvent this problem when studying phonetic contrasts by trying to use contrasts where the acoustic difference between the standard and the deviant was identical between conditions, and having a phonetic contrast between the standard and the deviant in one of the conditions (Sharma et al., 1993; Maiste et al., 1995). As it turned out these studies did not find an increase in MMN response in the phonetic contrast conditions. Maiste et al., (1995) suggested that the mismatch response to phonetic contrasts might be much smaller than the response to acoustic contrasts.

The introduction of variation among standards may be another more promising technique to differentiate between acoustic and phonetic effects that contribute to the MMN.

As already discussed above (1.4.2.1.2 MMN & Memory), the MMN is elicited when an activated representation of invariance is changed or violated, independent of the complexity of these invariances. MMN elicitation therefore tolerates some range of stimulus variation, as long as some rule or pattern is shared by the standard (e.g. Gomes, Ritter, & Vaughan, 1995; Huottilainen, Ilmoniemi, Lavikainen, Tiitinen, Alho et al., 1993; Winkler, Paavilainen, Alho, Reinikainen, Sams, & Näätänen, 1990; Phillips et al., 2000a). This makes the MMN an excellent tool for studying language and natural speech sounds, since language sounds obviously vary within and between each other. The first study with speech sounds using acoustically varying standard stimuli in the same experimental block was done by Aulanko et al. (1993), where they varied the F0 of their standard and deviant stimuli (/ba/-/ga/), while keeping all other stimulus aspects constant. By introducing naturally varying speech sounds in the oddball paradigm, some acoustic variability is brought in. But still sharing the same key phonetic and phonological features the processing system is forced to map the incoming variances onto a more abstract pattern of representation, grouping them perceptually. When the waveforms were averaged over the acoustically different stimuli of each phonetic category, a mismatch response indeed occurred (Aulanko, et al., 1993; Shestakova et al., 2002; Phillips et al., 1995; 2000a, 2000b). It was concluded, that the effect was purely phonetic, since the acoustically different tokens were grouped together into a more abstract bound representation of a speech sound.

This approach, introducing variation and using natural speech sounds, as well as using the identity MMN approach will be used in our studies, as they will help to control for physical stimulus properties in a better way, and will help to differentiate between MMN responses to purely acoustic or phonetic processes.

The idea behind using the MMN to study language processes, in particular phonetic sound representations is the process generating the MMN as described above. When a sound is presented as a deviant it elicits an activation pattern in the auditory system, which carries specific stimulus information of sensory features such as phonological features towards higher systems. This means a sound percept is created and forms a so-called surface representation that corresponds to features extracted from the acoustic signal, such as phonological features. The standard sound binds features into a perceptible auditory representation – the central sound representation – activating information of the underlying mental representation stored in long-term memory and may activate phonological representations in the mental lexicon during language processes. The change by the deviant reflects – besides the acoustic change – the comparison of the surface representation with the underlying mental representation activated by the standard, thus reflecting the output of a comparator that signals a mismatch between the stored memory trace and the physical

features of a deviant stimulus (Ritter et al., 1995). With this comparative process, incoming sound features of phonological stimuli can be compared against phonological features stored in mental representation in the mental lexicon.

This logic of the MMN, together with the reviewed evidence from MMN research suggests that this change detection response may be used as a pre-attentive measure of the early stages of spoken word processing. More specifically, the MMN data propose that the MMN may reflect not only acoustic processing of auditory material, but also the phonetic/phonological analysis of spoken language processing, and at least the initial stages of lexical processing. Therefore it is a brilliant instrument for understanding how speech sounds and phonological contrasts are stored in mental representations. For this reason the MMN will be used in all of our studies as the main method.

Last but not least, it is assumed that the activation of this automatic auditory change detection mechanism reflected by the MNN may also trigger the switching of attention to potentially important events in the unattended auditory environment (Näätänen et al., 2007). Evidence for the MMN being related to automatic attention shift is provided by the fact that the MMN is often followed by a relatively sharp peaking *fronto-central P3a*. This could indicate the occurrence of a short attentional shift or orienting to a deviance (Lyytinen et al., 1992; for review see Friedman, Cycowicz, & Gaeta, 2001) in the unattended sound sequence.

The next section will introduce this attention-related P3 component, which so often follows the MMN and seems relevant to some of our sound contrasts.

1.4.2.2 Stimulus Updating: The P300

A late positive waveform of the event-related brain potential to visual and auditory stimuli, the P3 or P300 (or ‘late positive component’ (LPC)) appears to be closely associated with the cognitive processes of perception, selective attention, and memory. The classical P300 deflection emerges in a time locked record as a positivity typically peaking at 300 to 400 ms after stimulus presentation. Timing of this component may range extensively, starting from 250 ms and extending to 900 ms, with amplitude varying from a minimum of 5 μV to a usual limit of 20 μV for auditory and visual evoked potentials. This large deflection was first described by Sutton and colleagues (1965) and is perhaps the most-studied (language independent) ERP component in investigations of selective attention and information processing. The P300 is often elicited using the ‘oddball paradigm’, in which occasional relevant (‘target’) stimuli have to be detected in a train of frequent irrelevant ‘nontarget’ or

standard stimuli. The subject is instructed to respond mentally or physically to the target stimulus and not to respond otherwise (*active* oddball paradigm). Such paradigms reliably yield P300 responses with a parietocentral scalp distribution to target as compared to standard stimuli irrespective of stimulus (visual, auditory, somatosensory) or response (button-press, mentally counting) modality (for review see Linden, 2005; Polich, 2007, 2004; Coles, Smid, Scheffers & Otten, 1995). The P300 is measured by assessing its amplitude and latency. P300 scalp distribution is defined as the amplitude change over the midline electrodes (Fz, Cz, Pz), which typically increases in magnitude from the frontal to parietal electrode sites (Johnson, 1993; cited in Polich, 2007).

The terms ‘P3a’ and ‘P3b’ indicate a distinction between two sub-components, where the P3b (or classical P300) is elicited by task-relevant deviant stimuli that are attended to. Whereas, a rare distinct tone presented in a series of frequent tones without a task can produce a positive wave that has an earlier deflection with a frontocentral scalp distribution and typically shorter latencies (Squires, Squires, & Hillyard, 1975; Knight, Scabini, Woods, & Clayworth, 1989; Katayama & Polich, 1998). This P3a is thought to reflect frontal lobe functioning (Friedman & Simpson, 1994; Knight, 1984; cited in Polich, 2004). Several ERP components have been reported that appear to be related to the P3a, one of which is elicited by rare, unexpected (‘novel’) stimuli inserted randomly as a third type into target / standard sequences or is elicited by targets that are irrelevant for the task but are unique and more salient than the trains of repeated standard and target stimuli. This potential has been called the ‘novelty P300’ (Courchesne, Hillyard, & Galambos, 1975). It is observed across modalities (e.g. Fabiani, Friedman, & Cheng, 1998; cited in Polich, 2007) and populations (e.g. Friedman & Simpson, 1994; cited in Polich, 2007). Studies support the interpretation that the P3a and the novelty P300 are variations of the same ERP component and generation system (Simons, Graham, Miles, & Chen, 2001; Spencer, Dien, & Donchin, 1999; Polich & Comerchero, 2003; Combs & Polich, 2006; cited in Polich, 2007, 2004) that vary in scalp topography as a function of attentional and task demands. Simons et al. (2001) conducted a study, in which they elicited the P3a and the novelty-P300 with prototypical procedures in order to address their relationship. Their data suggest that these two components are very similar, and potentially the same and concluded that together with others (Knight, 1996; Tsuchiya, Yamaguchi, & Kobayashi, 2000; Lagopoulos, Gordon, Barhamali, Lim, Li et al., 1998; Spencer, Dien, & Donchin, 2001; cited in Simons et al., 2001) their results are ‘more consistent with an identity hypothesis’ (p. 216), meaning that both sub-components are in fact identical.

Thus, the P300 phenomenon is thought to be composed of several parts that reflect an information-processing cascade when attentional and memory mechanisms are engaged (Polich, 2007). The parietal P3b is proposed to reflect match / mismatch with a consciously

maintained working memory trace (match between the stimulus and voluntary maintained attentional trace), thus reflecting the operation of memory storage operations. Whereas, the frontal P3a reflects a passive comparator, a reflection of the attentional switch produced from the mismatch between a stimulus presented and passively formed neural trace (Nätäänen, 1990). It has also been interpreted as a neural correlate of the orienting response (Friedman et al., 2001; Soltani & Knight, 2000) and comprises an early attention process stemming from a frontal working memory system (Polich, 2007). It is evident that the P3b and the P3a are dependant on the integrity of different brain structures, since Knight et al. (1989) has shown that only the P3a is influenced by lesions of the frontal cortex.

The differentiation between P3b and P3a subcomponents has begun to clarify the interaction between initial and subsequent P300 processes, but the knowledge of the P300 mechanism remains vague. The fact that there is more than one P300 with each having different latencies, topographies and amplitude characteristics, more research is necessary to develop a theory¹⁶ about the different P300 aspects.

¹⁶ Evidence has accumulated describing the P300 as a cognitive routine, in which a stimulus is evaluated: e.g. the 'context-updating' model (Donchin & Coles, 1988; McCarthy & Donchin, 1981). In this model, the P300 indexes brain activities underlying revision of the mental representation stimulated by incoming stimuli (Donchin, 1981). In other words, memory and schemata are consistently being updated after unexpected incoming information has been evaluated. After early processing of sensory information, a memory comparison process driven by attention assesses the representation of the previous event in working memory. The current stimulus of the oddball sequence is compared to the previous stimulus. If no stimulus feature change is detected, the current mental scheme of the stimulus context is preserved, however, if a new stimulus is detected, the system engages attentional processes to 'update' the neural representation of the context that is associated with the P300. This 'context-updating' hypothesis may reflect relatively strong initial target stimulus processing more related to the P3a and weakens as the repeated standard stimuli occur to elicit the P3b. However, the updating of memory requires more than just the binding of a new pattern. Human memories are possibly represented in a complex bound trace with many different recollections applied to an extensive neuronal network (Metcalf, 1993; cited in Picton, 1992).

Furthermore, the generation of the P300 linked to attention and memory-related operations might be caused by brain mechanisms engaged to inhibit irrelevant brain activation (*inhibition hypothesis*). Such an inhibitory process is required for adequate storage of memory contents in long-term memory. The suggestion is that the P300 and its underlying sub-processes could reflect rapid neural inhibition of online activity to assist interaction of stimulus / task information from frontal to temporo-parietal locations. Early attention processes from a frontal working memory change elicits the P3a. The attention-driven stimulus signal is then transmitted to temporal and parietal sites (P3b) to sharpen memory operations. The initial need to enhance attention during the stimulus detection relative to the contents of working memory would be signalled in the P300 (Soltani & Knight, 2000; e.g. Knight, 1997; cited in Polich, 2007). This theory is supported by neurophysiologic results (e.g. Friedman et al., 2001; Nieuwenhuis, Aston-Jones, & Cohen, 2005) and experimental findings (e.g. Birbaumer & Elbert, 1988; Birbaumer, Elbert, Canavan, & Rockstroh, 1990; cited in Polich, 2007).

An alternative theory is the *Context-Closure Theory* (Verleger, 1988) reflecting the concept that the P300 mirrors activity of memory trace remodelling post-detection of a target stimulus, as interplay of activation and deactivation. The difference to the context-updating theory is that the P3 is evoked by expected target incidents rather than by unexpected events, making the closure of a perceptual period. However, further work has proven only weak support for the view that the P300 is evoked by expected rather than unexpected events (e.g. Verleger & Hopmann, 1989; cited in Verleger, Jaskowski, Wascher, 2005).

In the next section the relationship between the MMN – as main ERP component of this work – and the P3a will be discussed briefly. The P3a could be elicited in some of our sound contrasts, especially those that seem more salient.

1.4.2.3 MMN & P3a

The distinction between the two components is predominantly based upon the fact that the MMN is relatively unaffected by either the relevance of the stimulus to any task or the amount of attention that the subject is paying to the stimuli (Näätänen, et al. 1978). The MMN is not dependant on attention (Alho, Sainio, Sajaniemi, Reinikainen, Sams, & Näätänen, 1991; cited in Ritter et al., 1995), whereas the P3a is. These two ERP components are proposed to index different processes along the course of distraction (for review see Escera, Alho, Schröger & Winkler, 2000). The MMN marks instances when the preattentive system detects sounds which may carry previously unavailable information that require conscious processing, being associated with the initiation of involuntary attention switch towards a change – a ‘call’ to attention (Giard et al., 1990). The P3a component in turn, is regarded as a reflexion of attention switching itself. The MMN is elicited by incoming sounds that violate some previously invariant characteristics of an auditory stimulus sequence. The process generating the MMN is initiated by a mismatch between the incoming sound and some memory record representing the regularities of the immediate history of auditory stimulation. One can conclude that incoming sounds are extensively analyzed and their regularities – including abstract ones are detected even when these sounds fall outside the focus of attention. In fact, the MMN process can be regarded as an information filter: it is activated only when a sound could not be ‘predicted’ from the preceding stimulus sequence, i.e. carrying new information. Therefore, one might suspect that the MMN could be involved in the passive (bottom-up) redirecting of attention (e.g. Alho et al., 1994; Schröger, 1996; cited in Escera et al., 2000). On this basis, Näätänen (1990) proposed that the process reflected by the MMN initiated redirecting of focused attention to infrequent changes in the auditory background. The existence of frontal MMN generators supports the assumed role of the MMN in calling for focal attention (Öhman, 1979; cited in Escera et al., 2000). Moreover, the MMN often precedes the P3a component, even to small changes in the to-be-ignored stimulus sequences, an ERP sign of attention switching. This response might indicate the actual orienting of attention to an MMN-eliciting sound change (Sams et al., 1985a), which could be of relevance for our sound contrasts tested in this work.

In certain circumstances, the information from the MMN system is transmitted to a frontal lobe mechanism that might make the event available to consciousness and behavioral control, thus activating the P3a. For example, if the event is sufficiently deviant, the MMN is followed by the P3a (Lyytinen et al., 1992). Evidence for this sequential neural course comes from results of experiments in which the processing of task-relevant stimuli delivered in one ear is disrupted by task-irrelevant stimuli in the second ear (e.g. Gaeta, Friedman, Ritter, & Cheng, 2001; Escera, Alho, Winkler & Näätänen, 1998; cited in Friedman et al., 2001). Although the MMN does not seem to reflect the involuntary catch of attention or orienting, the P3a does, suggesting a dissociation of processes between the MMN and the P3a.

As mentioned above, these two components also have been described as part of a specific pattern of brain responses, the ‘distraction potential’ (DP), showing a characteristic triphasic shape, starting with a negative wave followed by a positive one, and ending with a final phase of a sustained negative potential (*three-stage sequential model of distraction*; Escera and Corral, 2007). These waveforms provide a neurophysiological index of what is considered the three main processes involved in involuntary attention control (Escera et al., 2000): the mechanism of attention capture, associated with the MMN and/or with the N100, the orientation of attention, associated with the P3a, and the reorientation of attention toward main task performance after momentary distraction, associated with the re-orienting negativity (RON; Schröger & Wolff, 1998a; cited in Escera et al., 2000). Studies reviewed in Escera & Corral (2007) support the idea that the MMN generating process is drawing attention toward unexpected violations in the implicit regularity of the acoustic environment. This process would operate automatically. In contrast, it is proposed that the fronto-central P3a reflects the evaluation of contextual novelty of unexpected sounds, which signals the reconfiguration of a neural network involved in updating goal-directed action selection (Barceló, Escera, Corral, & Periáñez, 2006; cited in Escera & Corral, 2007).

However, Horváth, Winkler and Bendixen (2008) found dissociations between the ERP components hypothesized to reflect the three-stages of distraction. They investigated distraction-related ERP responses elicited by sound change following repetitive micro-sequences (‘local deviants’). Their P3a component appeared to be sensitive to the first repetition of a task-irrelevant feature after a change in that feature, whereas the N100/MMN was not. Additionally, the P3a amplitude increased with increasing number of repetitions of a task-irrelevant feature preceding a change in that feature, whereas this had no effect on the RON. They concluded that the ERP responses associated with subsequent stages of distraction are not strongly linked proposing that the P3a rather reflects some, possible higher-level event-detection mechanism, than attention switch itself.

Further research is necessary to uncover the relationship between the MMN and the P3a, as well as the specific functional roles of each of these ERP components. The P300 was

introduced since this component may occur in some of our speech contrasts. However, the P3 is not used for specific predictions about our sound contrasts rather it might co-occur when attention comes into play.

The N100 and the RON components will not be further discussed in detail, since this would go beyond the scope of this work. However the following literature on these topics give a good overview and insight (*N100*: Näätänen & Picton, 1987; Giard, Perrin, Echallier, Thevenet, Froment, & Pernier, 1994; Escera et al., 1998; Hari, Kaila, Katila, Tuomisto, & Varpula, 1982; Näätänen & Winker, 1999; *RON*: Schröger & Wolff, 1998a, 1998b; Escera, Yago, & Alho, 2001; Berti & Schröger, 2001; Schröger, Giard, & Wolff, 2000).

2 General Methods

2.1 Participants

All subjects who participated in the studies described in this dissertation were students of the University of Konstanz. They were all right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971; score of (R-L)/(R+L) greater than 0.08 were included in the data analysis), native German speakers with normal hearing and no history of neurological or psychiatric disorders. All subjects of Experiment 3-5 passed a standard hearing screening evaluating individual hearing thresholds for 200-8000 Hz (Audio Console Vers. 2.4.8; Oscillo Hearing Instruments, Inmedico, Denmark). All subjects signed an informed consent form and were paid for their participation or received course credit points. Details about the number of subjects of each study, their age (mean age, age range; standard-error of mean 'SEM') and gender distribution will be described in each separate experiment.

2.2 Stimuli

The standard and deviant stimuli of all six experiments were speech sounds embedded into words, nonwords or syllables.

Experiment 1 used three minimal pairs of German nouns varying exclusively in their mid-vowels [e], [ø] and [o] of the first syllable, with additional three matched nonwords. Experiment 2 varied three naturally spoken CV-syllables with plosive onsets. Experiment 3-6 manipulated VCV-syllables such as [edi], [egi], [eni], [ezi], [ebi], and [emi] which are all nonwords in German.

All sounds were checked for their frequencies by evaluating their log values from cumulative frequency counts weighted for word frequency based on token counts of the CELEX corpus (Baayen, Piepenbrock, & Gulikers, 1995). Distributional frequencies of the speech sounds (phonotactic probabilities), meaning the sequential arrangement of phonemes in the syllables and words were also controlled for the first critical consonant-vowel ([CV]; Experiments 1 & 2), consonant-vowel-consonant ([CVC]) syllables (Experiment 1), as well as for the first critical vowel-consonant ([VC]) and vowel-consonant-vowel ([VCV]) syllables (Experiments 3-6). It has been shown (Bonte et al., 2005) that the distributional probabilities of phoneme clusters influence the processing of speech as well as the amplitude characteristics of the MMN (Picton et al., 2000). These probabilities were determined by log-values of the frequency counts (token counts) of phoneme sequences weighed for word frequency of the German Celex database (Baayen et al., 1995).

All stimuli were controlled for length and equalized for intensity (root-mean-square power; RMS) to control acoustic properties as best possible, since the MMN is highly sensitive to even minor acoustic differences between the standard and deviant stimuli.

Furthermore, pitch and intensity measures were controlled for best possible¹⁷. The similarities and differences between the stimuli are presented in each study individually, illustrating the acoustic differences and similarities between the contrasts of each stimulus set.

All stimuli were normalized to have the same peak sound energy. For the analysis and processing of the stimuli the PRAAT software was used (Boersma & Weenink, 2007).

The stimuli were presented with a fixed inter-stimulus interval (ISI) of 500 ms binaurally via headphones (except for Experiment 2 with an ISI of 800 ms). By using different variants of the same stimuli and by presenting them randomly, acoustic variability was introduced to simulate more natural speech perception and thus forcing the processing system to map the incoming acoustic signals onto more abstract representations (Eulitz & Lahiri, 2004; Jacobsen et al., 2004; Phillips et al., 2000a, 2000b).

2.3 Experimental Procedure

In all five MMN studies each participant was presented with a passive oddball paradigm while his/her EEG was recorded. The behavioral Experiment 6 used an active oddball paradigm while their response times were measured. All Experiments used a classical oddball design except for Experiment 2, which had a roving-standard design (see 1.4.2.1.4 ‘Variations of the Oddball Paradigm’). Each stimulus served as the standard and as the deviant, which made it possible to compare responses to the same stimuli of the standards and deviants to guarantee the same physical differences between the sounds when calculating the difference wave (iMMN; Pulvermüller & Shtyrov, 2006a; Phillips et al., 2000a, 2000b). During the study, 85% standards and 15% deviants per sound contrast and block were presented (Experiment 2 used 80% standards and 20% deviants per sound contrast). A pseudo-random stimulus sequence was created so that there were at least 5 ± 2 standard stimuli between any two deviants (mean 5.5). In each experimental session the sound categories were combined in all possible pairs of syllables, words and nonwords. The order of blocks was counterbalanced across subjects.

¹⁷ Acknowledgments

I want to thank Henning Reetz very much for his help with the description of the stimuli in providing input and scripts for the pitch and intensity contours, as well as the sound frequency values for all experiments. I would also like to thank for the technical support given by A. Bobrov, G. Salagan, N. Bekemeier and R. Grutschnig.

2.4 Data Acquisition and Analysis for the MMN Experiments (1-5)

Participants were seated in an electrically and acoustically shielded chamber. They were instructed to watch a silent movie of their choice and to ignore the stimulus sounds which were presented with a neurobehavioral presentation software (Presentation Version 13) via headphones.

The EEG was continuously recorded (TMS International, Type Porti-S/64 for Experiments 1 & 2; Brain Vision, Brain Amp DC; 2 x 32 channels for Experiments 3-5) from 64 electrode positions (EasyCap, Montage M10 to 10% System) against Cz as reference. All electrodes were average referenced. The AC signals were recorded and sampled at 512 Hz for Experiments 1 and 2, and sampled at 500 Hz for Experiments 3-5. Interelectrode impedances were kept below 5 k Ω . Vertical and horizontal eye movements (EOGs) were co-registered bipolar with an additional electrode located on the forehead in order to correct the EEG raw data for eye movements using the algorithm implemented in the Brain Electric Source Analysis software (BESA; MEGIS Software GmbH, Gräfelfing, Germany). During the study participants were asked to sit quietly and avoid excessive eye movements.

Further off-line data processing included a digital band-pass filter set to 1-30 Hz (0.03-30 Hz for Experiment 1) and a standardization from 64 channels to 81 channels. Event-related potentials (ERPs) were gained by averaging epochs, which started 100 ms before the stimulus onset and ended 700 ms thereafter; the time interval from -100 to 0 was used as baseline (Experiment 1 started 200 ms before word/nonword onset and ended 700 ms thereafter; the time interval from -200 to 0 was used as baseline). Recording epochs with eye movements were corrected, and the maximum voltage criterion was kept below 100 μ V at any EEG channel epochs, for other recording artifacts to be rejected before averaging. Channels with exceeding noise levels were interpolated. A maximum of seven electrodes were interpolated with a maximum of two side by side. Standards immediately following deviants were not included in the analysis. All responses were re-referenced offline against right and left mastoids for further analysis (Schröger, 1998).

Following Eulitz and Lahiri (2004) the MMN as difference waveform was obtained by subtracting the response to the standard from that to the deviant stimulus. The so-called same-stimulus differences (identity MMN; Pulvermüller & Shtyrov, 2006a; Phillips et al., 2000a; 2000b) were calculated by using the recordings of two corresponding blocks. This method provides a reliable measure to avoid confounds by variation in ERP morphology that may result from stimulus differences per se (see also 1.4.2.1.5 ‘The MMN as Difference Response controlling the Effect of Physical Stimulus Differences’).

For statistical analyses (using SPSS Statistics 15.0), the individual MMN mean amplitude and the individual peak latency at the Fz electrode (re-referenced against linked

mastoids) was used as dependant variables for each experiment. The peak latency was determined as the latency of the most negative point in the difference waveform of each individual subject for each condition. The mean amplitude was measured over a 40 ms time window centered at the peak latency of the averaged difference waveform, separately for each individual subject in each condition.

Further analyses were calculated for Experiments 2-5 taking the MMN mean amplitude at the frontal electrode (Fz) and the central electrode (Cz) of condition-related peaks. The mean amplitude was measured over a 40 ms time window centered at the peak latency of the averaged difference waveform, but separately for each grand-averaged condition. These analyses however, will only be displayed in the Appendices under 5.2.2 Statistical Tables.

Details about the parameters for the statistical analysis, as well as the statistical analysis will be described in each experimental section under '*Data Acquisition and Analysis*' separately for each experiment.

3 Experiments on Place-and Manner-of-Articulation Features

3.1 Processing and Representation of Place Features

3.1.1 Experiment 1: Underspecification of the Place Feature [CORONAL] in Vowels¹⁸

One issue in neurolinguistics is the question of the precise nature of mental representations involved in speech processing which has been widely discussed (Poeppel, Idsardi, & van Wassenhove, 2008; Stevens, 2002; for review see Klatt, 1998; McQueen, 2005; see also main introduction in Chapter 1). Different psycholinguistic models have very different views on the nature of mental representations. Some have been discussed in the introduction. The *Featurally Underspecified Lexicon Model* (FUL; Lahiri & Reetz, 2002, 2010), is one example of abstract approaches to lexical representations. It claims that there is only one single underlying representation for each morpheme and predictable phonological information is not stored. The assumption of abstract and sparse information in lexical representation can account for many variances occurring in natural speech especially asymmetries occurring in assimilation processes. This tolerance results from the three-way matching algorithm of the FUL model (match – non-mismatch – mismatch) where not all matches are perfect and not all non-matches lead to a conflict. Underspecified representations are the prerequisite for tolerating surface variants.

This sort of argument was used by Connine, Blasko, and Titone (1993) although they were not suggesting underspecification. In their study initial and medial consonants of words were changed in minimal features (e.g. initial phoneme change: *teacher* ~ **deacher* or medial phoneme change: *manner* ~ **mammer*) to produce nonwords. These items were able to activate lexical representations. Only models assuming underspecified representations such as FUL would predict this pattern of results. All other types of models discussed in the introduction, like surface-form approaches (Bybee, 2001), or context dependent and experience-related accounts (Gaskell & Marslen-Wilson, 1997, 1999; Norris, 1994) would not expect lexical activations in this type of experiment.

In this study we distanced ourselves from any possible contextual influences, rather [e], [ø], and [o] focuses on vowel alterations. These alterations were not served by assimilations – they are distinct vowels in German. In an earlier study, Eulitz and Lahiri (2004) examined isolated vowels also free of context. However, isolated vowel sounds can

¹⁸ This study was presented at the annual meeting of the Cognitive Neuroscience Conference (CNS) 2008: Cornell, S.A., Lahiri, A. and Eulitz, C., (2008). ‘What you encode is not necessarily what you store’: An MMN study of phonological processing. *Supplement of the Journal of Cognitive Neuroscience* 290.

form lexical items (e.g. English [ai] ‘I’, Dutch [y] ‘you - honorific’, Bengali [o] ‘he/she’), they are few in number. To examine whether such vowel alterations behave in the same way in disyllabic complex structures, we compared them in pairs of words and nonwords.

The possibility of asymmetric activation was examined in this study by using the MMN. It has been proposed that the sound percept created by the deviant stimulus corresponds to the surface representation, formed by phonological features that are extracted from the acoustic signal. The series of standards presented before the deviant, form a central sound representation where the information structure is closer to the format used for long term memory (e.g. Cowan et al., 1993; Näätänen et al., 1993), i.e. close to the so called underlying representation in the mental lexicon. When standard and deviant come together in the MMN design the surface representation of the deviant is compared to the underlying representation of the standard. Thus, the MMN as change detection response reflects – besides the acoustic change – the comparison between the surface and the underlying representation (Eulitz & Lahiri, 2004).

The nature of mental representations of phonological features has been studied by means of the MMN for vowels in isolation (Eulitz & Lahiri, 2004). Eulitz and Lahiri (2004) used the same three vowels [e], [ø], and [o] as in our present study, where the phonological features in the signal are [DORSAL] for [o], and [CORONAL] for [e] and [ø]. Importantly, the acoustic formant frequency differences of [e] and [ø] are equidistant to the difference of [o] and [ø] (see Figure 7 & Methods 3.1.1.1). However the representational format is different, in that [e] and [ø] have the same phonological features, which are assumed to be underspecified and [o] being different from the other two having a specified place of articulation. Following the FUL-model (Lahiri & Reetz, 2002, 2010), when [e] and [ø] are presented as deviants after a sequence of /o/ sounds where [DORSAL] is pre-activated, a phonological conflict occurs, since [CORONAL] is extracted from the deviant and mapped onto a specified mental representation. In contrast, if the standards and deviants are reversed, i.e. when /ø/ and /e/ as standards precede [o] as the deviant, then no conflict is expected. Eulitz and Lahiri’s (2004) crucial result was the following: The same acoustic contrasts triggered asymmetric MMNs when they were reversed as standard and deviant in the coronal – dorsal contrast. Earlier latency and higher amplitude MMN values were found for the coronal vowel deviants [e] and [ø], when [DORSAL] was pre-activated by the standard /o/, and smaller values for the reversed case, with the underspecified coronal vowels serving as standards. However, the coronal vowel contrast [e] and [ø] did not differ in amplitude or latency independent of whether they served as standard or deviant (see also Figure 3). Eulitz and Lahiri (2004) interpreted these differential MMN asymmetries for similar acoustic/phonetic differences between pairs of vowels in isolation as reflecting the fact that the brain refers to underspecified phonological representations.

The aim of the current study was to examine the nature of mental representations for vowels in the mental lexicon. First, medial vowel alterations in word and nonword¹⁹ pairs were examined. It was of interest whether vowels in medial positions in complex linguistic structures reliably evoke MMN effects similar to vowels in isolation. The second issue concerns the possible interactive effects of the lexical status of the stimuli with the phonological level of feature conflicts.

With respect to the first goal, phonological models assuming full specification (Bybee, 2001; Gaskell et al., 1995; Gaskell & Marslen-Wilson, 1997, 1999; Norris, 1994; Norris & McQueen, 2008) expect to find equal MMNs between all contrasting vowels placed in words and nonwords, independent of the direction of presentation of the standard and deviant. As mentioned before, only models presuming underspecification of phonological features predict an asymmetry of MMNs within the reversal of vowel contrasts, presented as standard and deviant (for summary see Table 2). As in Eulitz and Lahiri (2004), we also expected higher MMN amplitudes for our vowel contrast if the place feature [CORONAL] that is extracted from the deviant [ø] maps onto the place feature [DORSAL] in the mental representation created by the standard /o/, which is assumed to be a conflict. In the reversed comparison when the feature [DORSAL] from the deviant [o] maps onto the underspecified representation of the standard /ø/ with no information about place of articulation, a non-conflict occurs and smaller MMNs are expected. Thus, in the acoustic contrast [o]_{/ø/} ~ [ø]_{/o/} (labeling see Table 2), an asymmetric MMN pattern would be expected. For our control contrast, [ø]_{/e/} and [e]_{/ø/}, we do not predict asymmetric MMNs, since the feature [CORONAL] is extracted from the acoustic signal of the deviant in both vowels and does not conflict with the underlying representation that is created by the corresponding standard.

With respect to the second issue, i.e. the possible interaction of phonology and lexicality, a number of studies have shown that the MMN is influenced by the lexical status (Endrass et al., 2004; Jacobsen et al., 2004; Pettigrew et al., 2004a; Pulvermüller et al., 2004; Pulvermüller et al., 2001; Shtyrov et al., 2008; Shtyrov & Pulvermüller, 2002b; Yasin, 2007). Words seem to evoke a larger MMN-amplitude in a nonword context (created by standards) compared to nonwords in a word context. Following these results, the possible impact of the lexical status of stimulus material on the MMN asymmetries between our vowel contrasts might be an issue. To control for possible influences of lexicality as well as to get an idea whether or not it is equally efficient to use nonwords when studying phonological effects in MMN-designs, the word blocks of the present study were paralleled by matched nonword blocks. Importantly, our study did not directly compare word-nonword

¹⁹ The term ‘nonword’ is used as ‘pseudowords’ in this work, with legal phonotactic sequences without semantic content.

sequences in one block as it usually has been done in the literature concerning lexical effects on the MMN. Rather, we used separate word and nonword blocks.

Table 2: Predictions about the amount of feature conflict for the vowel contrasts of models assuming full specification and underspecification (i.e. no specification of the coronal place of articulation in the underlying representation), correspondingly.

Experimental conditions [Deviant]/Standard/	Place features in the surface representation (extracted by the deviant)	Place features in the lexical representation (activated by the standard)	Full specification hypotheses	Underspecification hypotheses
[ø] _{/o/}	[CORONAL]	[DORSAL]	conflict	conflict with [DORSAL]
[o] _{/ø/}	[DORSAL]	[]	conflict	non-conflict with []
[ø] _{/e/}	[CORONAL]	[]	conflict	non-conflict with []
[e] _{/ø/}	[CORONAL]	[]	conflict	non-conflict with []

The shaded area highlights the only critical condition [ø]_{/o/} for the model assuming underspecification. Note, that the predictions are the same for the word and the nonword conditions. The term ‘conflict’ refers to ‘mismatch’.

3.1.1.1 Methods

3.1.1.1.1 Participants

A total of 24 students participated in this EEG study. The data of eight participants had to be excluded because of excessive eye movement artifacts, cardiac or muscle artifacts and therefore were not part of the statistical analyses. The remaining 16 participants (age range 20-28; mean age 23 (SEM .63); 50% women) went into the final analysis.

3.1.1.1.2 Stimuli

The standard and deviant stimuli were three minimal pairs of German nouns (they had to be the German dative plural forms to produce pairs; subjects were instructed correspondingly) varying exclusively in their mid-vowels [e], [ø] and [o] of the first syllable ([me:rən] ‘oceans’, [mø:rən] ‘carrots’, [mo:rən] ‘swamps’) and three matched nonwords ([me:kən], [mø:kən], [mo:kən]) (Figure 6). Due to constraints on the lexicon, triplets matching in all necessary aspects were hard to find. Consequently, there were some lexical inflectional ambiguities in the words. For instance [mo:rən] has two possible lexical sources. On the one

hand, it could be the dative or accusative form *Mooren* of the word ‘swamp’, as well as the nominative plural form *Mohren* of the word ‘Moor’. Further, [mø:rən] can be the nominative or dative plural of the same word *Möhren*, meaning ‘carrot’. Additionally [me:rən] can be the dative plural of *Meer* (‘ocean’) or the infinitive of the verb form *mehren* (‘to increase’). To ensure that we had exactly the same phonological context such that for each word there was a parallel nonword, we had to compromise on the inflectional ambiguities of the words and opted for more lexical complexity. The subjects were instructed that they would hear inflected words. No ambiguity occurred for the nonwords.

Words and nonwords were produced by a male speaker (30 years). The items were recorded multiple times and three variants per item were selected with comparable pitch and formant frequencies. The nonwords were phonotactically legal and had no meaning in German. The stress was on the first syllable for all stimuli.

The acoustic structure of all stimuli was fairly similar, starting with a nasal [m] followed by varying mid-vowels [e], [ø], and [o], and a second syllable that was [-ren] for the words, and [-ken] for the nonwords. For the present study, the critical differences were on the vowels of the first syllable that differed acoustically mainly in F2 and F3 formant frequencies (Figure 7). The coronal vowel pair [ø] and [e] varied in F2 and F3, whereas [o] and [ø] only varied in F2. However the distance between [ø] ~ [e] and [ø] ~ [o] were similar and equidistant in the F2 dimension which is crucial for place of articulation information. Crucially, the acoustic differences were symmetric, but the feature distinctions were asymmetric. The sounds within one category were very close to each other. The F0 (97-101 Hz) was close to each other for all vowels, and the F1 was close to each other within each vowel category ([e] 287-314 Hz; [ø] 305-336 Hz; [o] 496-535 Hz). Furthermore, pitch and intensity measures were controlled for as best possible. Each set of words and nonword showed comparable variation in pitch and intensity (see Appendix Figure 5.1.1a & b).

The surface word forms did not differ drastically in their cumulated inflectional frequencies for their dative and nominative plurals, as well as their infinitive verb word forms. According to the Leipziger Wortschatz database (<http://wortschatz.uni-leipzig.de/abfrage/>) [me:rən], [mø:rən], and [mo:rən] were similar in their frequencies (see Table 3 for their log values).

Distributional frequencies of the speech sounds (phonotactic probabilities) were controlled for the first critical consonant-vowel ([CV]) and consonant-vowel-consonant ([CVC]) syllables (Table 3). The [CV] sequences were all similar to each other ([me] log-value 4.07; [mø] log-value 3.94; [mo] log-value 3.96), the [CVC] sequences showed the highest probability for [me:r] (log-value 3.89), followed by [mo:r] (log-value 2.68) and [mø:r] (log-value 1.23). Furthermore, the individual vowel frequencies in medial position (C[V]C) were determined in the same way as the phoneme probabilities (see also General

Methods 2.2 Stimuli). The [o] had the highest frequency (log-value 5.99) followed by [e] (log-value 5.23), and [ø] (log-value 4.09) (see Table 3).

To avoid differential effects of early acoustic cues, the critical first syllables [mø:], [mo:] and [me:] were all separated from their second syllables and cut to the same length (250 ms). In order to achieve complete acoustic similarity in the standard-deviant contrasts across stimulus pairs, we then cross-spliced the first syllables with the second syllables [-rən] and [-kən]. For example, the original [-rən] from [mø: -rən] was attached to a different [mø:-] originating from either [mø: -kən] or a different [mø: -rən]. This technique allowed us to circumvent a primarily acoustic explanation for differences between word and nonword blocks, since the MMN is highly sensitive to even minor acoustic differences between the standard and deviant stimuli.

The stimuli (words and nonwords) of 588 ms duration were presented every 1088 ms with a fixed ISI of 500 ms. By using six different variants of the same stimuli and by presenting them randomly, acoustic variability was introduced to simulate more natural speech perception.

Table 3: Log-values of the frequency counts of phonemes and phoneme sequences in the studied words

word stimuli	CVCVC*	C[V]C	[CV]	[CVC]
[me:rən]	2.17	5.32	4.07	3.89
[mø:rən]	2.20	4.09	3.94	1.23
[mo:rən]	1.46	5.99	3.96	2.68

Frequency counts were weighed for word frequency and are based on token counts of the CELEX corpus. [...] indicates the phonemes for which counts are given in the table; [V] = vowel; [C] = consonant; Note, that the frequency counts for [CV] is the same for words and nonwords. C[V]C stands for the individual frequency of each vowel following [m] and preceding [r] or [k] (e.g. m[o]r, m[o]k, m[e]r, m[e]k, etc.). *The inflectional frequencies (log-values from cumulative frequency counts weighted for word frequency) based for the dative and nominative plurals (CVCVC) were taken from the Leipziger Wortschatz database (<http://wortschatz.uni-leipzig.de/abfrage/>).

Oscillogram

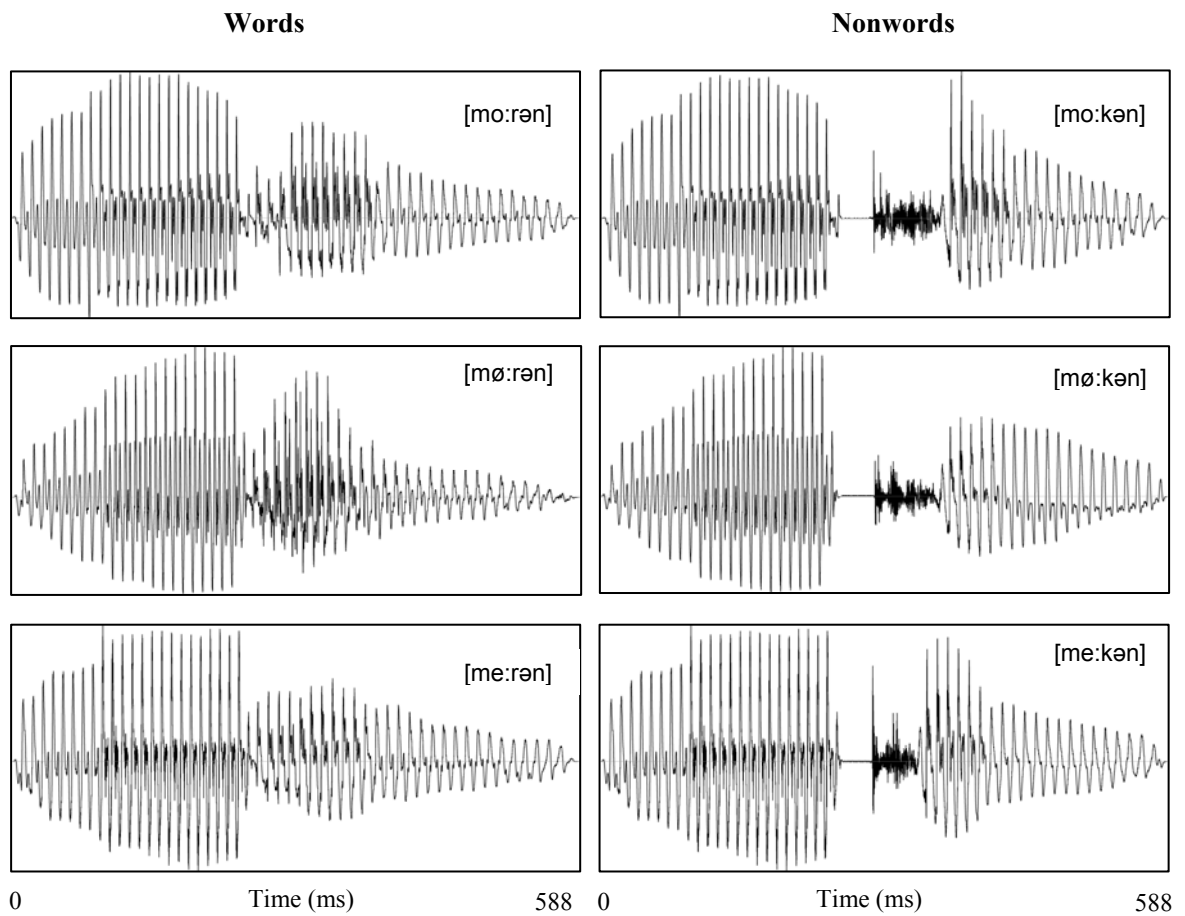


Figure 6: Oscillograms of the acoustic stimuli used in the experiment: words: [mo:rən], [mø:rən], [me:rən] and nonwords: [mo:kən], [mø:kən], [me:kən]. The standard and deviant stimuli were maximally matched for their acoustic properties in all conditions.

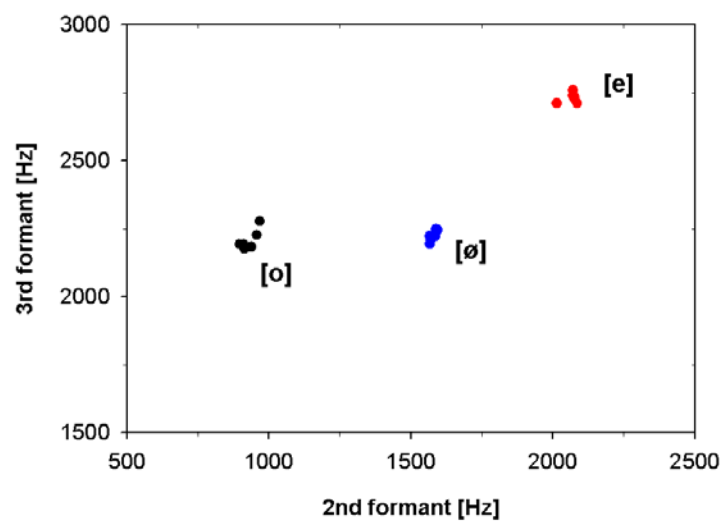


Figure 7: Acoustic characteristics of the three natural vowel categories used in the study. There are six exemplars of each vowel category. All vowels show their locations in the F2-F3 space. Note, that the distances between [e] - [ø] are equal to [o] - [ø] in F2.

3.1.1.1.3 *Experimental Procedure*

Each participant was presented with a passive oddball paradigm with 510 (85%) standards and 90 (15%) deviants per word and nonword category and block. In each experimental session, the three mid-vowel categories were combined in all possible pairs of words and nonwords, except for the [e] ~ [o] contrast, which we had to compromise, to ensure acceptable duration of the experiment. Each word and nonword served as a standard and deviant, resulting in eight blocks, four word blocks and four nonword blocks (Table 4).

Table 4: German words and nonwords used in the eight experimental conditions

	Conditions	Standard	Deviant
Words	[ø] _{/ø/}	[mo:rən]	[mø:rən]
	[o] _{/ø/}	[mø:rən]	[mo:rən]
	[ø] _{/e/}	[me:rən]	[mø:rən]
	[e] _{/ø/}	[mø:rən]	[me:rən]
Nonwords	[ø] _{/ø/}	[mo:kən]	[mø:kən]
	[o] _{/ø/}	[mø:kən]	[mo:kən]
	[ø] _{/e/}	[me:kən]	[mø:kən]
	[e] _{/ø/}	[mø:kən]	[me:kən]

The standard-deviant contrasts are identical within the reversal of vowel contrasts for words and nonwords. The shaded areas highlight the only conflicting condition [ø]_{/ø/}.

3.1.1.1.4 *Data Acquisition and Analysis*

The data were acquired and analyzed as described in the General Methods section (2.4 Data Acquisition and Analysis). Not mentioned before, the averaged MMN responses contained at least 80% accepted deviant trials in each condition for each participant.

The identity MMN was calculated by using the recordings of two corresponding blocks. For instance, the standard [mø:rən] (of the block with [mo:rən] as deviant) was subtracted from the [mø:rən]-deviant of the reversed block (with [mo:rən] as standard).

For statistical analysis, the mean amplitude and the peak latency was measured for each individual subject in each condition as the dependent variable. The mean amplitude was taken from the expected MMN latency window between 150-250 ms after change onset (vowel), thus in a window between 240-340 ms. Amplitude values of the difference waves were tested against zero in two-tailed *t*-tests separately for each condition to test for significant MMNs.

A three-way repeated-measures ANOVA for each mean amplitude value with the factors WORDNESS (word vs. nonword), PAIR OF INVERSION (showing an equalized acoustic change: non-conflicting pair [e]_{/ø/} ~ [ø]_{/e/} vs. conflicting pair [ø]_{/ø/} ~ [o]_{/ø/}) and DIRECTION OF ACOUSTIC CHANGE (of the F2 frequency between the deviant and standard: F2 falling: [ø]_{/e/} ~ [o]_{/ø/} vs. F2 rising: [e]_{/ø/} ~ [ø]_{/ø/}) was performed. Only significant main effects or interactions are reported. In the case of significant interactions paired *t*-tests were used for post-hoc testing. The Greenhouse-Geisser adjustment was used where appropriate and the corrected *p* values are reported together with the uncorrected degrees of freedom.

3.1.1.2 Results

To examine the possible interference of the lexical status with the phonological feature conflict we expected higher amplitudes for our words as compared to the nonwords. With regard to phonological effects we predicted an asymmetry for our conflicting PAIR OF INVERSION, in particular a higher MMN amplitude for the conflicting condition [ø]_{/ø/}.

A clear MMN in the expected time window of 240-340 ms of the grand average difference waveform was observed (Figure 8) in all experimental conditions (see Appendices for detailed statistical analyses (5.2.2 Statistical Tables) and figures (5.1 Figures) of the grand-average waveforms of standards, deviants and difference waveforms including their MMN maps in Figures 5.1.3a-d).

The overall ANOVA did not reveal a three-way interaction ($F(1, 15) = 0.27; p > .1$; n.s.), nor a main effect for WORDNESS ($F(1, 15) = 0.12; p > .1$; n.s.), indicating that in the present experiment no systematic impact of the lexical status on the pattern of MMN differences was seen. However, with respect to phonological effects, a significant two-way interaction was found for PAIR OF INVERSION and DIRECTION OF ACOUSTIC CHANGE ($F(1, 15) = 6.28; p < .024$) for the MMN amplitude at the frontal electrode position. Paired comparisons showed a significant difference in the PAIR OF INVERSION containing a phonological conflicting ([ø]_{/ø/}) than the non-conflicting condition ([o]_{/ø/}) ($F(1, 15) = 6.39; p < .023$), both pairs being of the same acoustic contrast. In comparison, the PAIR OF INVERSION with no phonological conflict in both conditions ([e]_{/ø/} ~ [ø]_{/e/}) did not differ significantly ($F(1, 15) = 1.20; p > .1$; n.s.). Paired comparisons between the DIRECTION OF ACOUSTIC CHANGE of F2 revealed only differences when F2 was rising (F2 rising: $F(1, 15) = 5.74; p < .03$; F2 falling: $F(1, 15) = 3.39; p > .1$; n.s.).

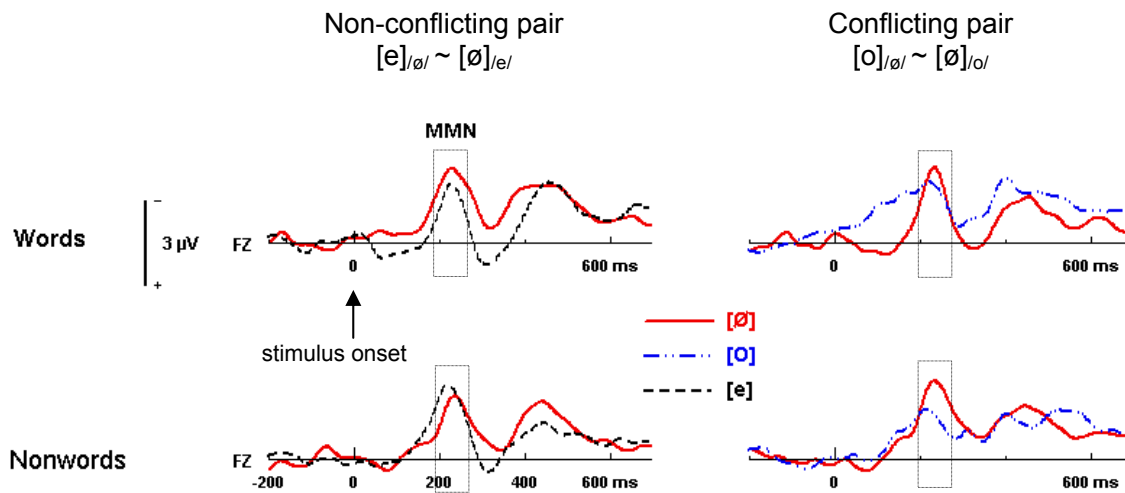


Figure 8: Grand-average difference waveforms of the frontal electrode position (Fz) for the non-conflicting ([ø]_{/e/} ~ [e]_{/ø/}) and the conflicting PAIR OF INVERSION ([ø]_{/o/} ~ [o]_{/ø/}) for words and nonwords. The color of the waveforms represents the deviant vowel, red for [ø], blue for [o], and black for [e]. Solid lines represent the conflicting conditions, dashed lines the non-conflicting conditions. The squared areas mark the MMN time window, showing the asymmetry for the conflicting conditions that is strongest for nonwords.

Hypothesis driven, we ran additional paired comparisons separately for the word and nonword conditions to discern whether or not the asymmetrical effects were equally robust in words and nonwords. A difference was found for the conflicting contrast [ø]_{/o/} ~ [o]_{/ø/} for nonwords only as a tendency ($t(15) = -2.06; p < .057$). Similarly, for words, there was again a tendency for a difference ($t(15) = -1.96; p < .069$). However, the non-conflicting contrasts for words as well as nonwords did not show any difference (all $t(15) > -1.0; p > .1; n.s.$). For detailed statistical analysis of the *MMN Mean Amplitudes* see Table 14a in the 5.2.2 Statistical Tables of the Appendices.

An additional analysis was conducted using the MMN peak amplitude of each individual subject for each condition. The *peak* amplitude values of the difference waves were again measured at the Fz electrode (re-referenced against linked mastoids) taken from the expected MMN latency window between 240-340 ms after stimulus onset. This analysis is more precise and can pronounce the findings from the mean amplitude analysis, although it is not used very often in the literature. Basically the same results were found as in the mean amplitude analysis, no three-way interaction ($F(1, 15) = 0.24; p > .1; n.s.$), as well as no main effect for WORDNESS ($F(1, 15) = 0.03; p > .1; n.s.$).

As for the phonological effects, the significant two-way interaction for PAIR OF INVERSION and DIRECTION OF ACOUSTIC CHANGE ($F(1, 15) = 5.21; p < .037$) was also found in this analysis. Paired comparisons again showed a significant difference between the PAIR OF INVERSION containing a phonological conflict than when a non-conflict occurred ($F(1,$

15) = 4.87; $p < .043$; Figure 8 & 9). In comparison, the PAIR OF INVERSION with no phonological conflict in both conditions ($[e]_{/o/} \sim [\emptyset]_{/e/}$) did not differ significantly ($F(1, 15) = 1.24$; $p > .1$; n.s.). Comparisons between the DIRECTION OF ACOUSTIC CHANGE of F2 did not reveal any differences this time (F2 falling: $F(1, 15) = 3.28$; F2 rising: $F(1, 15) = 3.7$; all p values $> .1$; n.s.).

This peak analysis showed more pronounced effects for the comparisons ran separately for the word and nonword conditions. This time the conflicting contrast $[\emptyset]_{/o/} \sim [o]_{/o/}$ was truly significant for nonwords – not only a tendency as in the mean amplitude analysis – ($t(15) = -2.44$; $p < .027$), showing a higher amplitude for the condition, where the deviant $[\emptyset]$ mapped onto the underlying representation of $/o/$. For words, the previous tendency in the mean amplitude analysis was ascertained here ($t(15) = -1.0$; $p > .1$). The non-conflicting contrasts for words as well as nonwords again did not show any differences (all $t(15) > -1.0$; $p > .1$; n.s.).

To summarize, this additional peak amplitude analysis revealed the same results as the mean amplitude analysis, but produced a more pronounced pattern for the conflicting contrast for the nonwords, this time differing clearly (for detailed statistical analysis of the *MMN Peak Amplitudes* see Table 14b of 5.2.2 Statistical Tables in the Appendices).

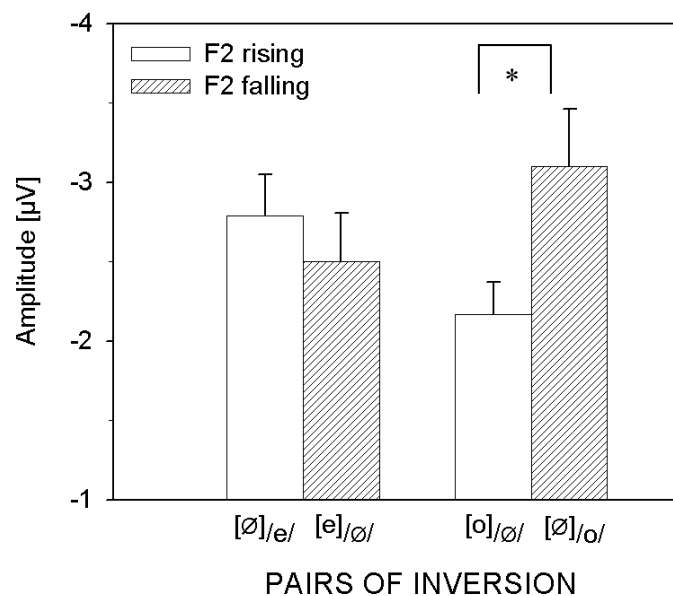


Figure 9: MMN peak amplitudes grand-averaged across words and nonwords for all conflicting and non-conflicting conditions. The striped bars indicate the DIRECTION OF ACOUSTIC CHANGE of the F2 formant falling, the solid bars represent F2 rising. Note the more pronounced amplitude difference for the PAIR OF INVERSION which includes the phonological conflict $[\emptyset]_{/o/}$. Error bars show the standard error of the mean (SEM).

The overall ANOVA of the MMN peak latency values did not reveal a three-way interaction ($F(1, 15) = 0.15; p > .1; n.s.$) either, only a main effect for WORDNESS ($F(1, 15) = 6.14; p < .026$). Words showed a slightly earlier peak (5 ms) than nonwords (for detailed statistical analysis of the *MMN Peak Latencies* see Table 14c of 5.2.2 Statistical Tables in the Appendices).

The absence of a significant interaction with WORDNESS in the MMN time window in either analysis was surprising. In order to ascertain the actual impact of lexicality onto the automatic processing of our stimuli, a further analysis was conducted using a later time window to test for the well-known N400 effect. This analysis was restricted to the mean amplitudes of the standard and deviant ERP data for the Fz in a time window between 350-450 ms which should well cover the N400 effect. This analysis revealed a clear N400 effect for words. They were more negative than nonwords. However, since this analysis goes beyond the scope of this work it was only included in the Appendix (5.2.1 Experiment 1: Standard/Deviant Analysis; Table 14d of 5.2.2 Statistical Tables in the Appendices), without any further discussion here.

3.1.1.3 Discussion of Experiment 1

The present MMN study was designed to examine whether the underspecification of the [CORONAL] place of articulation, which has already been demonstrated in isolated vowels (Eulitz & Lahiri, 2004), can be shown for linguistically more complex language processing conditions in medial position. When embedded in words or matched nonwords, the vowel contrasts examined, revealed supporting evidence for the underspecification of the [CORONAL] place of articulation in vowels. This conclusion is based on the fact that larger MMN amplitudes were found whenever a place-of-articulation conflict occurred compared to non-conflicting situations (Eulitz & Lahiri, 2004). According to the FUL-model (Lahiri & Reetz, 2002, in press), a phonological conflict occurs, when the deviant [ø] where the information about the coronal place of articulation is extracted from the signal and presented in a row of standard stimuli which have pre-activated a [DORSAL] place of articulation (an [o] in the present study). For the reversed condition when [o] is the deviant and /ø/ the standard, a non-conflict occurs, since [CORONAL] is underspecified. This current experiment found exactly this: larger amplitudes when the coronal deviant mapped onto a dorsal standard (conflict condition [ø]_{/o/}) compared to the reversed condition, where the dorsal deviant was compared to a standard coronal sound (non-conflict condition [o]_{/ø/}). This was pronounced for our nonwords.

Models assuming fully specified information (Bybee, 2001; Johnson, 1997; Gaskell & Marslen-Wilson, 1997; Norris, 1994; Norris & McQueen, 2008) would predict a similar phonological conflict for both directions of acoustic change.

Eulitz and Lahiri (2004) provided the first neurophysiological evidence for the notion of underspecification of the feature [CORONAL] in the mental lexicon when using isolated vowels. Our significant two-way interaction for PAIR OF INVERSION and DIRECTION OF ACOUSTIC CHANGE in the present study is clearly in line with the predictions of the FUL-model and do not support the idea of fully specified place-of-articulation information which would predict at best two main effects in the present experimental setup (for a summary of the predictions see also Table 2). Additionally, Gaskell's context-sensitive, experience-driven connectionist model (Gaskell, et al., 1995; Gaskell, 2003; Gaskell & Marslen-Wilson, 1997, 1999) cannot really explain our asymmetries since our stimuli are built in such a way that all varying vowels have the same preceding (/m/ preceded all vowels) and following contexts (for words /r/, for nonwords /k/), so no contextual information could have influenced our vowel asymmetries.

Additionally, the obtained MMN amplitude effects cannot be caused by well known factors of influence, like spectral, pitch and other acoustic differences (for review see Näätänen, 2001), as we studied MMN differences for acoustically equidistant, but phonologically asymmetric pairs of stimuli and used the identity MMN approach. The iMMN has been shown to be sensitive to a variety of linguistic parameters – especially important for this study – the sensitivity to feature specification of segments (Eulitz & Lahiri, 2004; for review see Pulvermüller & Shtyrov, 2006a).

With respect to the present results among the discussed parameters, a possible influence of phonotactic probabilities might be of relevance. Bonte et al. (2005) reported higher MMN amplitudes for nonwords with high phonotactic probability (*notsel*) as compared to the low probability condition (*notkel*). The distributional probabilities of C[V]C sequences in our experiment were highest for the clusters with [e] and lowest for [ø] stimuli. According to the findings of Bonte et al. (2005) the prediction for the present study is a higher MMN amplitude for [me:kən] and [me:rən] compared to all the [ø] and [o] conditions. However, the pattern of MMN differences in the present study was completely different. The conditions with the largest difference in the phonotactic probabilities showed no MMN differences whereas the conditions with a just moderate difference in phonotactic probabilities showed a reversed MMN difference, i.e. contrary to Bonte et al. (2005), larger MMN amplitudes for the low phonotactic probability conditions²⁰.

²⁰ Furthermore, the MMN findings cannot be explained by individual frequency effects of our stimuli. The individual sound frequencies of our vowel stimuli differ in that [ø] has the lowest frequency, with increasing frequency from [e] to [o] (see Table 3). Again, one could argue that a high frequency

Is there a possible common ground of both studies to explain the differential pattern of the reported MMN effects? From a phonological point of view, Bonte et al. (2005) used place- and manner-of-articulation differences in consonants to create a contrast between standard and deviant. At a featural level, the *notsel* - *notkel* difference includes a manner change (from [STRIDENT] to [PLOSIVE]) as well as a place change (from [CORONAL] to [DORSAL]). Given the underspecification of the [CORONAL] place of articulation (Lahiri & Retz, 2002) and the findings about the correlation of levels of feature conflicts with the MMN amplitude in the present study as well as in Eulitz & Lahiri's study (2004), the MMN asymmetries in Bonte's et al. results can be interpreted as a reflection of the underspecified [CORONAL] place of articulation in *notsel*. With respect to the contrasts of place features the stimuli of Bonte et al. (2005) follow the same systematic as described in Table 2 for the present experiment given that the place of articulation is [CORONAL] for the [ø] and [s] stimuli and [dorsal] for the [o] and [k] stimuli. In sum, in light of the results of the present experiment where the phonotactic probabilities were calculated as proposed by Bonte et al., (2005), we have no supporting evidence for the influence of phonotactic probabilities in the way as suggested by Bonte et al., (2005). Taken together, there is a hint that the MMN effects in Bonte et al. (2005) might have been driven by differential conflict levels between segmental phonological features in the same way as in the present study and as in Eulitz and Lahiri's study (2004). If this is true, Bonte et al. (2005) would be one of the first studies (together with Walter & Haquard, 2004), which suggest a generalization of the original results for vowels to consonantal speech sounds. Further research is, however, necessary to study these aspects in a controlled fashion.

A second goal of the study was to get an impression about possible inferences of other levels of linguistic processing with the perception of phonological contrasts and its impact on the MMN effects. Therefore, word blocks and matched nonword blocks were recorded. The stimuli were designed in a way that the final syllable (offset of the crucial vowel) could complete a word or a nonword. It is important to note that up until the time of the first syllable (CV-syllable), the processing system has not yet received full information about the lexical status. The deviation point of the lexical status comes after the crucial vowel. It turned out that the asymmetry for the conflict vs. non-conflict conditions was more clearly pronounced for our nonword conditions. Although the interaction with the repeated measures factor WORDNESS was not significant, it was seen in the post-hoc tests, as well as

deviant would elicit a higher MMN response; however, in our results the MMN amplitudes show a different pattern. Following these vowel frequencies, the deviant [o] would have to have the highest amplitude compared to the deviant [ø], however the reversed was found in the MMNs. Also [e] and [ø] would have to show a difference which was not the case. These results show a pattern, which cannot be explained by frequency effects.

reflected in the ERP waveforms. We interpreted this as a result of a possible superimposition of additional factors of influence, which could not be fully balanced when using words. As mentioned in the methods section our words could not be fully controlled. The words in isolation could have a different meaning or case for some but not all words or could have been interpreted as a verb. Although we made sure in the instructions that the ambiguity was minimized and the words were repeated many times in the blocks, this kind of superimposition might have taken place. As we said before, well-grounded conclusions along this line of argumentation are not licensed by the corresponding interactions. However, the slightly less pronounced phonological effects for the word conditions gave us a hint that lexicality may interfere even in the relatively early time window of the MMN component. It might be argued that through the strong repetition effect created by the standards the lexical ambiguities are being tremendously diluted. To be on the safe side for follow-up experiments, we decided to use nonwords as stimuli whenever possible. Note that the present tests for WORDNESS effects cannot be compared to the previous experiments in this field (see Pulvermüller & Styrov, 2006; Jacobsen et al., 2004) where the difference in the lexical status was introduced between standards and deviants. Here we focused the experimental design on possible inferences on other levels of linguistic processing with the perception of phonological contrasts and the possible impact on the MMN.

When working with nonwords it might be argued that they do not tap into lexical representations at all, rather that the nonword standards do not activate mental representations. However in order to decide whether the incoming speech sounds is a word or nonword the lexicon indeed must be activated, thus activating the same features independent of the lexical status. Connine et al. (1993) argued that words and similar-sounding nonwords are not treated equivalently. In her studies she finds priming effects for derived words (nonwords that are minimally changed in one or two features) that are reduced to those for words. However, she concludes that the lexical item activated in memory is simply the best hypothesis available for the acoustic input, thus tapping into the same lexical items independent of the lexical status.

In sum, the present study extends the neurobiological findings about the underspecification of the [CORONAL] place of articulation, which has been demonstrated in isolated vowels (Eulitz & Lahiri, 2004), to more complex language processing situations, such as words and nonwords. The present results also show that rather phonological contrasts at the segmental level were primarily driving the MMN effects and not alternative factors of influence from higher levels of linguistic processing such as phonotactic probabilities. The reported pattern of MMN differences support the notion that mental representations of phonological place features for vowels are not a one-to-one relation between the acoustic speech signal and their mental representations, but may be more

abstract compared to variabilities in the surface realization of lexical forms, as assumed by theories suggesting the storage of all phonological details.

This reported MMN asymmetry between conditions for the same acoustic contrasts strongly suggests that the brain refers to more abstract underspecified phonological representations in the mental lexicon during speech perception and that underspecification may be an important principle of the functional organization of the mental lexicon. Additional research is, however, necessary to further generalize this notion, for instance from vowels to other speech sounds or to other featural dimensions. As a first step, Experiment 2 examines stop consonants in syllable initial position contrasting [LABIAL], [CORONAL] and [DORSAL].

3.1.2 Experiment 2: Underspecification of the Place Feature [CORONAL] in Stop Consonants²¹

The present MMN study was aimed at teasing apart effects to phonetic contrasts by systematically manipulating the place of articulation, in order to find further evidence for the generalization of the underspecified feature [CORONAL] as a basic principle in the functional organization of the mental lexicon. One goal was to expand the findings of phonological asymmetries for place-of-articulation features with vowels (Eulitz & Lahiri, 2004; Experiment 1) to other classes of speech sounds, here using stop consonants in a syllable initial position. There is a large body of MMN research using consonants (for review see Näätänen, 2001; Pulvermüller and Shtyrov, 2006a), many of which also contrasted place of articulation in stops (Walter & Hacquard, 2004; Maiste et al., 1995; Sharma et al., 1993; Dehaene-Lamberg, 1997; Rivera-Gaxiola et al., 2000; Sharfer et al., 2004; Diesch & Luce, 1997; Pettigrew et al., 2004a, 2004b). However, many of these studies focused on other issues such as categorical perception, and not so much on fine-grained differences in consonants to answer specific questions about phonological feature information and possible differences in representations.

Eulitz and Lahiri (2004) and Experiment 1 studied three vowels contrasting [CORONAL] and [DORSAL]. It was found that the same acoustic contrasts triggered asymmetric MMNs when they were reversed as standard and deviant in the coronal-dorsal contrast. Earlier and higher MMN amplitude values were found for the phonological conflict, when [CORONAL] was extracted by the deviant and mapped onto [DORSAL] in the underlying representation, than for the reversed non-conflicting case. These MMN asymmetries for similar acoustic/phonetic differences between pairs of vowels were discussed as neurobiological evidence for the brain referring to underspecified phonological representations.

In line with this research, this experiment uses the three consonantal sounds [b], [d], and [g] in CV-syllables, all being voiced stops denoted as [PLOSIVE] with differing place of articulation, the first being [LABIAL], then [CORONAL], and the last being [DORSAL] (in Table 5, the second column lists the surface features of the consonants in square brackets listed in the first column). With these sounds all possible place contrasts can be created, such as [LABIAL] ~ [CORONAL], [LABIAL] ~ [DORSAL] and [CORONAL] ~ [DORSAL]. Models, like FUL

²¹ Acknowledgment

This second experiment (plosives in initial position) was conducted by Christian Scharinger in the framework of his master's thesis. He was in charge of constructing and controlling the speech stimuli, collecting all the ERP data and further processing them off-line. The data presented in this dissertation were re-analysed and parameterized. I would like to thank Christian Scharinger for these data and his work.

(Lahiri & Reetz, 2002, 2010), presuming underspecification of phonological features, predict an asymmetry of MMNs within the reversal of place contrasts, presented as standard and deviant (for summary see Table 5), if [CORONAL] is part of the contrast. Therefore larger MMN amplitudes are expected in conditions where a feature conflict occurs as compared to those where no conflict occurs.

The predictions for the three place contrasts, or rather the six experimental conditions were the following: For the contrast [LABIAL] ~ [CORONAL], a conflict occurs only in one direction: for the conflict condition [d]_{/b/}, when the deviant [d] extracts the [CORONAL] feature from the signal and maps onto the previously activated specified feature [LABIAL] in the mental representation from the standard /b/. In the reversed condition [b]_{/d/} a non-conflicting situation occurs, when the feature [LABIAL] from the deviant [b] maps onto the pre-activated representation of the standard /d/ which contains no information about the place of articulation. Consequently, the MMN amplitudes between these acoustically equalized conditions [d]_{/b/} (conflict) and [b]_{/d/} (non-conflict) should differ with the conflicting condition showing the larger amplitude, i.e. an asymmetric MMN pattern would be expected. The [DORSAL] ~ [CORONAL] contrast is expected to show a similar asymmetrical MMN pattern as the [LABIAL] ~ [CORONAL] contrast, since again a conflicting condition ([d]_{/g/}) is compared to the reversed condition ([g]_{/d/}) where no conflict is expected. The third place contrast [LABIAL] - [DORSAL] operates as control condition, since this contrast is predicted to be symmetric. Both features are assumed to be fully specified for their place of articulation, so that a conflict is created in both directions of change, in [b]_{/g/} as well as in [g]_{/b/} which should not result in different MMN amplitudes.

Mental lexicon models assuming full specification like surface-form approaches (Bybee, 2001; Johnson, 1997), or experience-related accounts (Gaskell et al., 1995; Gaskell & Marslen-Wilson, 1997; Gaskell, 2003; Norris, 1994; Norris & McQueen, 2008) would not expect asymmetric activations between all contrasting stops with varying place features. Keep in mind that these models depend on experience or on contextual information. However, our stimuli are built in such a way that context does not help, since they are in an initial position and are followed by the same vowels.

Table 5: Predictions about the amount of feature conflict for the CV-contrasts used in the six experimental conditions

Experimental conditions [Deviant]/Standard/	Place features in the surface representation (extracted by the deviant)	Place features in the lexical representation (activated by the standard)	Full specification hypotheses	Underspecification hypotheses
[d] _{/b/}	[CORONAL]	[LABIAL]	conflict	conflict with [LABIAL]
[b] _{/d/}	[LABIAL]	[]	conflict	non-conflict with []
[g] _{/b/}	[DORSAL]	[LABIAL]	conflict	conflict with [LABIAL]
[b] _{/g/}	[LABIAL]	[DORSAL]	conflict	conflict with [DORSAL]
[g] _{/d/}	[DORSAL]	[]	conflict	non-conflict with []
[d] _{/g/}	[CORONAL]	[DORSAL]	conflict	conflict with [DORSAL]

All stimuli were matched for their acoustic properties. The standard-deviant contrasts are identical in each pair of conditions. The shaded area highlights the control contrast, the only bidirectional conflicting pair of inversion ([g]_{/b/} ~ [b]_{/g/}). [] = underspecified feature.

3.1.2.1 Methods

3.1.2.1.1 Participants

A total of 20 subjects took part in this EEG study. The data of six subjects were excluded due to excessive eye movement (EOG), cardiac or muscle artifacts after visual inspection. Therefore, the subsequent analyses are based on the remaining 14 participants (age range 19-27; mean age 23.87 (SEM .67), 50 % female).

3.1.2.1.2 Stimuli

The experimental stimuli used as standard and deviants were three naturally spoken CV-syllables, such as [bɛ], [dɛ] and [gɛ] (Figure 10). They were created from the initial [CV] syllable of a CVCCVC-syllable structure embedded in a fixed German sentence context ('Er hat [b/d/g]etten gesagt.' – 'he said [b/d/g]etten'; [b]etten means 'bed', [d]etten and [g]etten have no meaning in German). The sentences were spoken by a phonetically trained male speaker with normal articulation, intonation, loudness and speaking rate.

All sentences were recorded and digitized at a sampling rate of 44.100 Hz, (Mono, 16 Bit) with the sound application Software CoolEdit (V.5.4). The positioning of the CV-syllables in the sentences assured a stable and fairly equal articulation of the target syllables

and additionally guaranteed a natural pronunciation of the short vowel [ɛ]. The short vowel (e.g. [ɛ] like in ‘*bed*’) was chosen to avoid possible lexical effects on the MMN (e.g. Jacobsen et al., 2004, Pulvermüller et al., 2001). This vowel is only found in closed CVC-syllables in German (Wiese, 2000). Therefore, the CV-syllables with this vowel should be quite independent of the place-of-articulation change, always being perceived as nonwords and thus, free of lexical effects on the MMN.

A total of 40 target CV-syllables were extracted from the sentences and further processed in PRAAT (Boersma & Weenink, 2007). Four variants of each CV-syllable were selected with similar fundamental frequencies (F₀; [bɛ]: $\mu_{F0} = 93$ Hz, [dɛ]: $\mu_{F0} = 92$ Hz, & [gɛ]: $\mu_{F0} = 92$ Hz).

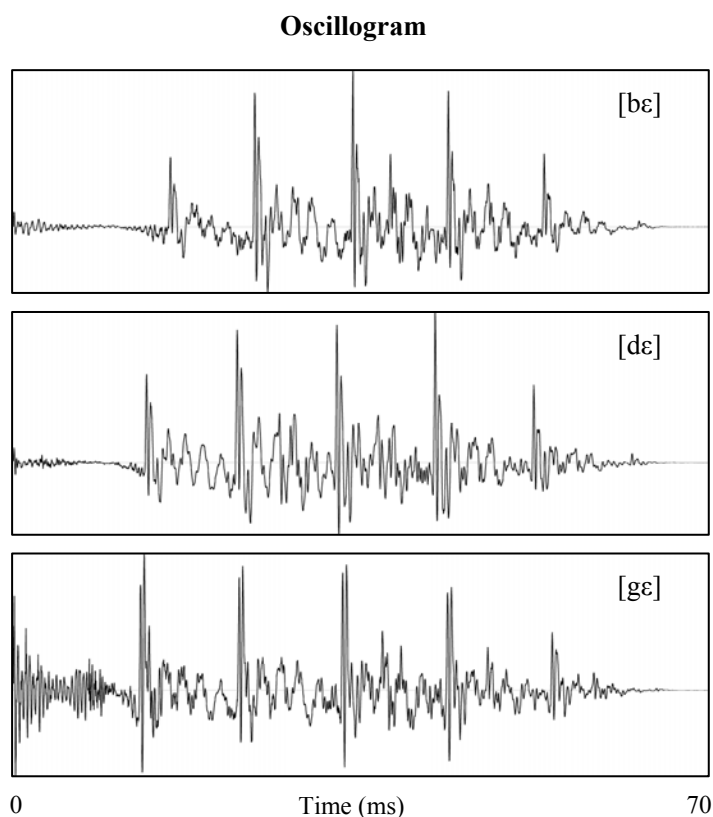


Figure 10: Oscillograms of the acoustic stimuli used in the experiment: [bɛ], [dɛ], [gɛ]. The standard and deviant stimuli were maximally matched for their acoustic properties in all conditions.

In order to avoid timing differences in the MMN, the bursts between all stimuli were adjusted by cutting a minimal mid-part within the burst (between the release and before the voice onset time), resulting in a mean burst length across all stimuli (mean burst length 15 ± 2 ms). The burst of the dorsal stop underwent the most reduction since it is characterized by the longest burst. Contrastingly, the labial and coronal stops were hardly altered. The length of the plosive sounds was therefore held constant, which resulted in similar acoustic parameters and still allowed a minimal natural variation between the stimuli of each

category. All CV-syllables were equally long (70 ms). They had equalized sound energy (RMS) to avoid acoustic differences between the stimuli, thus, their measured sound pressure intensities were normalized to 62 dB_{SPL}, and inspected by measuring the sound intensities with an artificial ear (sound level meter type 2250; artificial ear type 4152). The intensities of the last 30 ms of each stimulus were reduced linearly (linear fade out; Adobe Audition 2.0) to ensure a smooth ending of the stimuli. Thereby, acoustic differences between the stimuli could be minimized, thus reducing the likelihood that the MMN would be determined by only one or several particular acoustic features. Each set of CV-syllables showed comparable variation in pitch and intensity (see Appendix Figure 5.1.2 for more detailed descriptions about these CV-syllables also see C. Scharinger, 2007).

Additionally, the plosives were checked for their word-initial frequencies (Table 6). All three plosive sounds are highly frequent but do not differ substantially from each other ([d] log-value 5.95, [b] log-value 5.34, [g] log-value 5.24). Another influencing factor on the MMN is the distributional probability of phoneme clusters (phonotactic probability; Bonte et al., 2005) of speech sounds. The [CV] sequences were fairly similar to each other ([dɛ] log-value 4.05; [bɛ] log-value 4.00), with [gɛ] having a slightly lower probability (log-value 3.83).

All target stimuli were judged by 4 independent German native speakers as natural sounding CV-syllables with clear place-of-articulation difference of the plosives. The stimuli were presented every 800 ms with a fixed inter-stimulus interval binaurally via headphones.

Table 6: Log-values of the frequency counts of phonemes and phoneme sequences of the CV-syllables:

CV-syllables	[C]V	[CV]
[bɛ]	5.34	4.00
[dɛ]	5.95	4.05
[gɛ]	5.24	3.82

Frequency counts were weighed for word frequency and are based on the CELEX corpus. C = consonant; V = vowel; e.g. [C]V in the second column gives the frequency counts for /b/, /d/, & /g/ in an initial position. [CV] in the third column gives the frequency counts for the sound sequences /bɛ/, /dɛ/, & /gɛ/.

3.1.2.1.3 *Experimental Procedure*

All participants were presented with a roving-standard paradigm. Due to experimental complexity of the design and time consideration, we opted for the roving-standard paradigm which is less time consuming for the participants and is assumed to produce the same results

as the classical oddball paradigm (Cowan et al. 1993; Winker et al., 1996; Ritter et al., 1995; Shestakova et al., 2002; Baldeweg et al., 2004; see General Methods section 1.4.2.1.4 ‘Variations of the oddball paradigm’). A ‘roving’ stimulus train consisted of a sequence of standard and deviant stimuli changing the plosive sounds from train to train. For example, the standard labial stop /b/ was followed by a deviant of a different plosive sound (e.g. coronal stop [d]) being followed by the next train of standards consisting of the same coronal stops. Thus, the first stimulus of a new stimulus train served as a deviant for the preceding train (sequence of 3-7 standards between any two deviants, ITI 800 ms). Across the entire roving-standard sequence the six possible standard-deviant contrasts [b]_{/d/}, [b]_{/g/}, [d]_{/b/}, [d]_{/g/}, [g]_{/b/}, [g]_{/d/} occurred equally frequent with 146 (± 2) deviants (20% deviants) and 4380 (± 60) standards (80% standards) per contrast (Table 5). Each session included four blocks that lasted 14 minutes each, with breaks as needed between the blocks.

A short forced-choice test was conducted after each experimental session in order to test the auditory discriminability of the stimuli categories. All four stimuli of each plosive category were presented to all subjects four times in a random order, resulting in 48 presented stimuli with an ISI of 2000 ms. The participants were asked to press the letter on a keyboard corresponding to the sound they had heard.

3.1.2.1.4 *Data Acquisition and Analysis*

For each participant, the averaged MMN responses contained at least 88 % accepted deviant trials in each condition. The MMN as a same difference waveform (iMMN) was obtained by subtracting the ERP response of the standard from that of the equivalent deviant stimulus, e.g. [bε] as deviant minus [bε] as standard. Based on visual inspection of the grand-average waveform, the maximum amplitude of the MMN difference waveform of every individual condition was used for quantification of the ERPs. The MMN was expected in the typical MMN latency range between 150-250 ms after stimulus onset.

For statistical analyses, the MMN mean amplitude and the peak latency at the Fz electrode were used as dependant variables. The statistical analyses were restricted to two separate repeated-measures ANOVAs for each mean amplitude value with two independent variables: the factors PAIR OF INVERSION (POI) showing an equalized acoustic change within the reversal of standard and deviant (POI 1: *Asymmetric Labial-Coronal* pair [b]_{/d/} ~ [d]_{/b/} vs. control condition *Symmetric Labial-Dorsal* pair [b]_{/g/} ~ [g]_{/b/}; POI 2: *Asymmetric Dorsal-Coronal* pair [g]_{/d/} ~ [d]_{/g/} vs. control condition *Symmetric Labial-Dorsal* pair [b]_{/g/} ~ [g]_{/b/}) and DIRECTION OF CHANGE (DOC) of the F2 frequency between the deviant and standard (*F2 Front-to-Back*: DOC 1: [d]_{/b/} ~ [g]_{/b/}; DOC 2: [g]_{/d/} ~ [g]_{/b/} vs. *F2 Back-to-Front*: DOC 1: [b]_{/d/} ~ [b]_{/g/}; DOC 2: [d]_{/g/} ~ [b]_{/g/}) was performed. Front-to-back refers to the direction of F2

movement in the course of stimulation; for example in the condition $[d]_{/b/}$, the standard $/b/$ is more frontal ([LABIAL]) compared to the deviant $[d]$, leading to a F2 movement from front to back. The back-to-front refers to the reversed case: the standard $/g/$ activates a place of articulation ([DORSAL]) which is further back than the deviant $[d]$ or $[b]$. Thus for the conditions $[d]_{/g/}$ & $[d]_{/g/}$ the movement of the F2 formant is from back to front.

Only significant main effects or interactions are reported. In the case of significant interactions, paired t -tests (t -Test, α level = .05) were used for post-hoc testing. Additionally amplitude values of the difference waves were tested against zero in two-tailed t -tests separately for each condition to test for significant MMNs. The Greenhouse-Geisser adjustment was used where appropriate and the corrected p values are reported together with the uncorrected degrees of freedom.

3.1.2.2 Results

Behavioral Results:

The identification and discrimination of the forced-choice perception test were quite good for all subjects. The $[b]$ and $[d]$ sounds were correctly identified to 99%, the $[g]$ sound to 85 %. The fourteen subjects who were included into the statistical analysis identified all three stop sounds unambiguously.

ERP Results:

With regards to phonological effects, we predicted an asymmetry in the MMNs for our PAIR OF INVERSION *Labial-Coronal* and *Dorsal-Coronal*, in particular a higher MMN amplitude for the conflicting conditions $[d]_{/b/}$ and $[d]_{/g/}$. For our PAIR OF INVERSION *Dorsal-Labial* we did not expect to find MMN differences, since bidirectional conflicts were expected.

MMN early time window:

Statistical MMNs (tested against zero) at the Fz electrode in the expected MMN time window between 150-250 ms were only found in four of the six experimental conditions: the *Labial-Coronal* contrast ($[b]_{/d/}$: $t(13) = -3.03$; $p = .010$; $[d]_{/b/}$: $t(13) = -3.62$; $p = .003$), the *Dorsal-Coronal* non-conflicting condition $[g]_{/d/}$ ($t(13) = -3.96$, $p = .002$) and the *Labial-Dorsal* condition $[b]_{/g/}$ ($t(13) = -2.71$, $p = .018$). For the *Dorsal-Coronal* conflict condition $[d]_{/g/}$ and for the *Labial-Dorsal* condition $[g]_{/b/}$ no significant MMNs were observed ($[d]_{/g/}$ $t(13) = -1.84$; $[g]_{/b/}$ $t(13) = -1.35$; $p > 0.1$; n.s.; see Figure 11 & 12A and Appendices for more detailed figures of the grand-average waveforms of standards, deviants and difference waveforms including their MMN maps in Figures 5.1.4a-c).

The two ANOVAs did not reveal any significant interactions of PAIR OF INVERSION X DIRECTION OF CHANGE (POI 1: *Labial-Coronal* vs. *Labial-Dorsal*: $F(1,13) = 0.94$; POI 2: *Dorsal-Coronal* vs. *Labial-Dorsal*: $F(1,13) = 0.11$; $p > .1$; n.s.), as well as no significant main effects, indicating that in this second experiment with plosives in initial position, no systematic differences were seen between the places of articulation in the MMN patterns (see Table 15a of 5.2.2 Statistical Tables in the Appendices for detailed statistical analysis). Although no significant effects were found for the two asymmetric contrasts *Labial-Coronal* and *Dorsal-Coronal* in the time frame of the MMN (150-250 ms), small differences were visible which can be seen in the Figure 11. In a later time window around 250-350 ms these effects became even more visible. Therefore another analysis was carried out for this later MMN time window, to pick up possible late MMN effects.

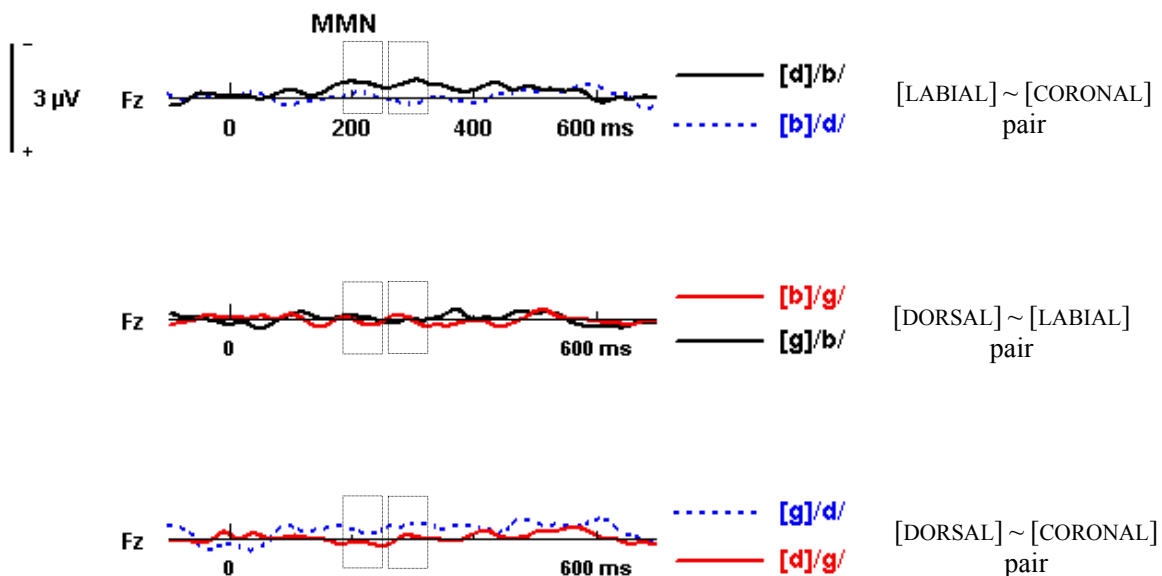


Figure 11: Grand-average difference waveforms of all six conditions (re-referenced against the right mastoid at Fz, the PAIR OF INVERSION *Labial-Coronal* ($[d]_{b/} \sim [b]_{d/}$), *Labial-Dorsal* ($[g]_{b/} \sim [b]_{g/}$), and *Coronal-Dorsal* ($[d]_{g/} \sim [g]_{d/}$). The non-conflicting conditions are shown in blue & dashed lines. Square areas mark the early (150-250 ms) and late MMN (250-350 ms) time windows at Fz.

MMN late time window:

As in the analysis of the early MMN effect, mean amplitude values were used as dependant variables of the difference waves at Fz in this later time window. A 40 ms time window for each individual subject for all six conditions generated the mean amplitude which was tested against zero in two-tailed t -tests separately for each condition, but this time for a later time window (250-350 ms). Repeatedly, the two separate repeated-measures ANOVAs for each

mean amplitude value was performed with the same independent variables, the two factors PAIR OF INVERSION (POI) and DIRECTION OF CHANGE (DOC).

This time only three of the six conditions showed significant MMNs ($[d]_{/b/}$: $t(13) = -3.03$; $p = .010$; the *Dorsal-Coronal* pair: $[d]_{/g/}$: $t(13) = -2.18$, $p = .049$ and $[g]_{/d/}$: $t(13) = -5.54$, $p = .000$), with $[g]_{/b/}$ showing a tendency to be significant ($t(13) = -2.01$; $p = .066$). The conditions $[b]_{/d/}$ and $[b]_{/g/}$ ($[b]_{/d/}$: $t(13) = -1.79$; $[b]_{/g/}$: $t(13) = -1.62$; $p > .1$; n.s.) did not show significant MMNs.

The two ANOVAs again did not exhibit any significant interactions of PAIR OF INVERSION X DIRECTION OF CHANGE (POI 1: *Labial-Coronal* vs. *Labial-Dorsal*: $F(1,13) = 1.16$; POI 2: *Dorsal-Coronal* vs. *Labial-Dorsal*: $F(1,13) = 0.34$; $p > .1$; n.s.), as well as no significant main effects, like in the earlier time window, indicating again no pattern of differences for the expected asymmetrical PAIRS OF INVERSION (Figure 11 & 12B; see Table 15a of 5.2.2 Statistical Tables in the Appendices for detailed statistical analysis).

The same analysis of the early and late MMN time window was calculated for condition-related mean amplitudes for Fz and Cz, in order to see whether there are any effects at all. This condition-related analysis was measured separately for each condition, resulting in six grand-average time windows for each condition. The analysis showed similar results as for the individual mean MMN amplitudes (see Table 15b & 15c of 5.2.2 Statistical Tables in the Appendices for detailed statistical analysis).

The individual peak latencies in the early and late MMN time window also did not show any differentiating effects (see Table 15d of 5.2.2 Statistical Tables in the Appendices for detailed statistical analysis).

Plosives: Place change

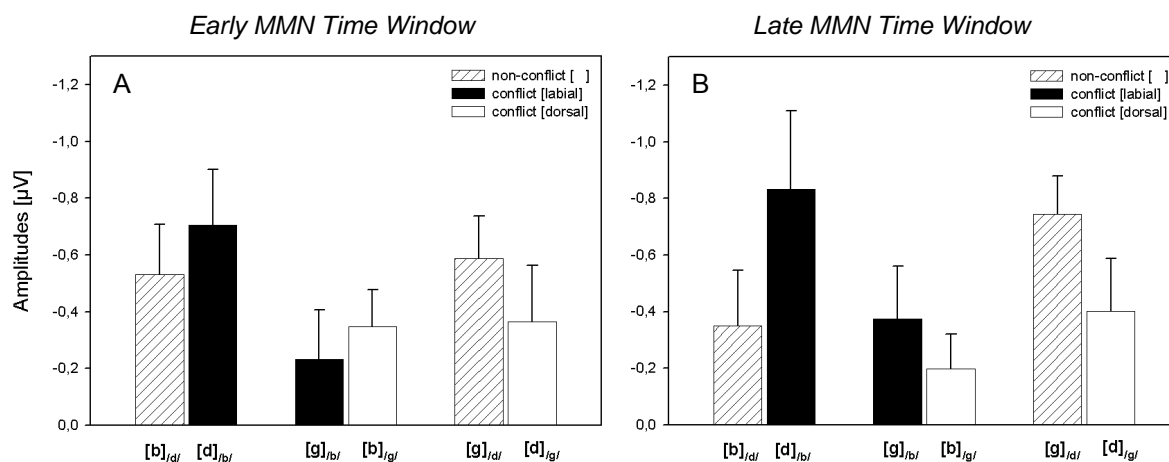


Figure 12: Mean amplitudes of the MMN difference waves for all six conditions at Fz, the asymmetric PAIR OF INVERSION *Labial-Coronal* [b]/d/ ~ [d]/b/, the symmetric pair *Labial-Dorsal* [g]/b/ ~ [b]/g/, and the asymmetric pair *Dorsal-Coronal* [g]/d/ ~ [d]/g/. Error markings on top of the bars represent the standard error of the mean (SEM). (A) shows the individual mean peak amplitudes of the early MMN time window (150-250 ms). (B) shows the individual mean peak amplitudes of the late MMN time window (250-350 ms).

3.1.2.3 Discussion of Experiment 2

This MMN experiment on place-of-articulation features with plosives in initial position was designed to examine the underspecification of the place feature [CORONAL] with stop consonants, in order to further generalize the previous findings of asymmetrical activations done with vowels (Eulitz et al., 2004; Experiment 1) and consonants (e.g. Friedrich et al., 2006; Friedrich et al., 2008; Walter & Haquard, 2004, Wheeldon & Waksler, 2004). Testing these speech sounds in syllable initial position, the stop contrasts under study did not show any significant interactions or main effects. In other words, the expected conflicting – non-conflicting asymmetries were not found. According to FUL (Lahiri & Reetz, 2002), a phonological conflict occurs, when the deviant [d] extracts information about the coronal place of articulation from the speech signal and is presented in a row of standards, such as /b/, which has pre-activated the [LABIAL] place of articulation, but not when the deviant is [b] and the standard is /d/. A conflict should induce a stronger neuronal activation pattern reflected in a higher MMN amplitude, whereas a non-conflict should exhibit a smaller amplitude. According to the predictions, this asymmetrical activation pattern should be the case for the *Coronal-Labial* and *Coronal-Dorsal* pairs, yet neither turned out to be significantly different. However, a tendency for the *Coronal-Labial* pair is visible (Fig. 11 & 12), especially in the later MMN time window, when the standard is a labial /b/ resulting in a

phonological conflict, the MMN appears to be larger. Nevertheless, this was not significant. The *Labial-Dorsal* pair as control condition also did not differ like the other two contrasts, but here this outcome was expected.

Altogether the MMNs in this experiment were very small. They were below 1 μ V, or not even significantly different from zero in some cases (early time window: [g]_{/b/}, [d]_{/g/}; late time window: [b]_{/d/}, [b]_{/g/}). This is in line with some other studies examining place of articulation with stops (e.g. Maiste et al., 1995; Shafer et al., 2004; Diesch & Luce, 1997; Pettigrew et al., 2004a, 2004b; Wunderlich & Cone-Wesson, 2001) finding very small or no MMNs. Diesch and Luce (1997) discuss the fact that plosive place-of-articulation contrasts seem to be more susceptible to interference by acoustic noise, i.e. are less discriminable than vowel contrasts. Furthermore, the discrimination of speech sounds encounter interference from acoustic noise, but also by noise caused within the nervous system. These speech sounds do not seem to show the same recency, modality, and suffix effects shown by vowels in short-term memory tasks that tap into auditory sensory memory (Cowder, 1971). This suggests that the auditory memory traces that are created by stops are weaker than the traces created by vowels (Diesch & Luce, 1997; Pettigrew et al., 2004b). This consonant-vowel difference holds for syllable-initial and syllable-final consonants (Cole, 1973). Thus, it is reasonable to suspect that the difference in the effortlessness of discrimination and retention between vowels and plosive stops might affect the amplitude of the mismatch responses, as it could be the case for the MMN responses in this experiment. The association between discrimination, retention in auditory memory, and mismatch response strength corresponds to similar findings for non-speech contrasts and vowels (Lang, Nyrke, Aaltonen, Paimo, & Näätänen, 1990; Imada, Hari, Loveless, & McEvoy, 1993; Winkler, Reinikainen, & Näätänen, 1993; Aaltonen et al., 1993; cited in Diesch & Luce, 1997). The lack of robust MMN responses of Diesch and Luce (1997), as well as in this current study are similar to the results of Pettigrew et al., (2004a), Pettigrew et al. (2004b) and Wunderlich and Cone-Wesson (2001), supporting the idea that MMN activation patterns to fine CV-syllables contrasts with initial stop consonants may be difficult to obtain. However, this finding remains surprising, given the richness of literature reporting robust MMNs to CV-syllables contrasting place of articulation with stop consonants (e.g. Kraus et al., 1992; 1993a; Aulanko et al., 1993; Bradlow, Kraus, Nicol, McGee, Cunningham et al., 1999; Uwer, Albrecht, & von Suchodoletz, 2002). Additionally, the speech stimuli were easily discriminable to all subjects at an attentional state, indicating that the poor responses were not caused by qualitatively poor speech stimuli. In particular, the studies by Dalebout and Stack (1999), and Dalebout and Fox (2000) further support the current finding that robust mismatch responses to these fine featural contrasts may not always be obtained although they are easily discriminated and identified in behavioral experiments.

One possible explanation for the diminished MMN amplitudes and the lack of differences between the altered places of articulation of the stop consonants could be rooted in the way the stimuli were built. The length of each stop consonant usually varies in its burst length giving information about its place of articulation (Halle, Hughes & Radley, 1957), causing different formant transitions and vowel onsets. Nevertheless in mismatch negativity experiments this could lead to timing differences in the MMN component. In order to avoid these latency effects on the MMN the bursts between all stimuli were adjusted by cutting a minimal mid-part within the burst (see 3.1.2.1 methods). This affected the dorsal stop the most since it is characterized by the longest burst, whereas there was hardly any need to change the burst of the labial and coronal stops. Manipulating the stimuli (by trying to reduce acoustic differences as best as possible) may have affected the MMN in a way that the place-of-articulation information was not strong enough to tap into representational differences. Making the experimental stimuli fairly similar (acoustically) and compromising for their naturalness might have erased the fine phonological differences.

Another consideration comes from the work of May, Tiitinen, Ilmoniemi, Nyman, Tylor and Näätänen (1999), suggesting that stimulation of the auditory system with pure tones in an oddball series might result in a pattern of adaptation and lateral inhibition across parts of the auditory cortex (Wunderlich & Cone-Wesson, 2001). This theory may help to explain the current findings for our speech contrasts. Spectrally complex stimuli, such as CV-syllables might result in a broad spread of activation over the auditory cortex (O'Leary, Andreasen, Hurtig, Hichwa, Watkins et al., 1996; Ottaviani, Girolamo, Brilia, Rossi, Giuda & Nardo, 1997). As a result, repeating presentation of the standard stimulus would set up a broad pattern of adaptation and lateral inhibition, affecting most of the areas, which are activated by the deviant sound. The amount of overlapping could be considerable and while the deviant would generate slightly different activity from that of the standard, it may not be sufficiently diverse to evoke a difference, which is measurable by the recording electrodes and visible in the MMN.

Additionally, the stimuli in the present study might have been too short, not giving enough information about the formant transitions for the MMN to capture the differences in place of articulation. Information contained in speech signals occurs on multiple time scales; for instance the time constant relevant for encoding formant transitions in stop consonants is approximately 20-40 ms (Poeppel, 2003). Poeppel (2003) proposed that these short temporal windows of integration are associated with gamma band activity. One should be able to measure slight differences in the gamma and theta spectral ranges. The stop duration in our stimuli was about 30 ms. One could argue that perhaps the lower frequencies of the MMN (1-30 Hz) did not capture our fine grained place differences and that the differences could be better captured if higher frequencies around 40 Hz of the gamma band were measured.

A more trivial explanation for the lack of differences between the different place contrasts may have been caused by the sentence context that the crucial consonants were embedded in. The critical consonants all followed a coronal [t] ('Er ha[t] XXXX gesagt'), which may easily have led to co-articulation effects. This means that all plosive sounds were influenced by the coronal sound [t], all being 'coronalized'. From a phonetic point of view, this is the most straightforward and easiest explanation.

It should be noted that the number of subjects in this study might not have been large enough, in order to gain significant differences, which were slightly visible for the *Coronal-Labial* contrast. In order to find significantly different phonological characteristics, we might need a larger group of subjects.

In summary, this present study on place-of-articulation contrasts with stop consonants in initial position was not appropriate for mapping phonological processes, nor acoustic differences of the F2 formant transitions. The information in the initial position was not sufficient to generate the expected differences, or to state anything about their representational nature. No MMN asymmetries were found between the different [CORONAL] ~ [LABIAL] and [CORONAL] ~ [DORSAL] contrasts for the early or late MMN time window, although this was expected assuming asymmetric representations in the mental lexicon in the framework of FUL (Lahiri & Reetz, 2002, 2010). Additionally, the other models assuming fully specified representations would have expected symmetric activation patterns of contrasting plosives. Not finding any reliable MMNs was not expected by either model. Therefore, this second experiment does not speak for either model. There might be other factors such as higher cognitive top-down processes influencing the MMN in a way that the fine phonological featural differences were disguised and overwritten. Plosive sounds that are discriminated on the basis of phonological differences cannot be relied on to evoke robust MMNs as many other studies have shown (Diesch & Luce, 1997; Pettigrew et al., 2004a, 2004b; Wunderlich & Cone-Wesson, 2001; Dalebout & Stack, 1999; Dalebout & Fox, 2000). Neural mechanisms, such as lateral inhibition and adaptation may prove to be critical in the interpretation of the MMN for these acoustically very similar plosive sounds.

3.2 Processing and Representation of Manner-of-Articulation Features

3.2.1 Experiment 3: Mental Representations of Place- and Manner-of-Articulation Features²²

In the previous experiment (Experiment 2) the fine structure of phonological place features was tested, in particular the underspecification of the feature [CORONAL] using stop consonants in initial position. The tested contrasts of initial stops were not very successful in tapping into mental representations. Hence, a further experiment was conducted involving stop contrasts in medial position in order to see whether a change in position might provide enough information to tap into these lexical representations.

One goal of this type of research is to expand the findings of phonological asymmetries for place-of-articulation features that already have been done with vowels (Eulitz & Lahiri, 2004; Experiment 1) and consonants in behavioral and neurobiological studies (Lahiri & Reetz, 2002; Wheeldon & Waksler, 2004; Friedrich et al, 2006; Friedrich et al., 2008; Walter & Haquard, 2004) to further speech sounds. We started out by investigating the stop consonants /d/ and /g/ in medial position, which contrast the features [CORONAL] and [DORSAL].

FUL assumes a universal system of phonetic contrasts, not only for place features, but also for other feature dimensions. Therefore, this third study was also aimed at teasing apart effects to phonetic contrasts by systematically manipulating two featural dimensions, place of articulation ([CORONAL] ~ [DORSAL]), and manner of articulation ([NASAL] ~ [STRIDENT]), in order to find further evidence for the generalization of underspecification as a basic principle of how the mental lexicon is organized. The manner feature contrast was manipulated in order to take a first step into another feature dimension. The question was raised whether manner-of-articulation features can be studied using the MMN as successfully as place features. So far manner features have not been looked at in a controlled fashion in an MMN design.

Vowels and consonants are made up of different featural properties. In particular, consonants differ in terms of place and manner of articulation, whereas vowels do not. Looking at the initial consonants of the words *deer*, *giraffe*, *zebra*, *nandu*, first of all, the sounds [d] and [g] are voiced stops denoted as [PLOSIVE] which differ in their place of articulation, the former being [CORONAL] and the latter [DORSAL]. Secondly, the fricative [z] shares the place of articulation of [d] and [n], where the latter has a [NASAL] manner of

²² Cornell, A.S., Lahiri, A. & Eulitz, C., (2009). 'MMN to phonetic contrasts along different feature dimensions'. *Frontiers in Human Neuroscience. Conference Abstract: MMN 09 Fifth Conference on Mismatch Negativity (MMN) and its Clinical and Scientific Applications in Budapest*, April. doi: s10.3389/conf.neuro.09.2009.05.079

articulation. The difference between [z] and [d] also rests on the manner of articulation where the former is a [STRIDENT] sound (in Table 7, the second column lists the surface features of the consonants in square brackets listed in the first column). It is assumed that there exist asymmetric relations between these featural properties.

Table 7: Predictions about the amount of feature conflict for the nonword stimuli used in the four experimental conditions

Exp. conditions [Deviant]/Standard/	Features in the surface representation (extracted by the deviant)		Features in the lexical representation (activated by the standard)		Full specification hypotheses	Underspecification hypotheses
	MoA	PoA	MoA	PoA		
[g] _{/d/}	[PLO]	[DOR]	[]	[]	conflict	non-conflict with []
[d] _{/g/}	[PLO]	[COR]	[]	[DOR]	conflict	conflict with [DOR]
[z] _{/n/}	[STR]	[COR]	[NAS]	[]	conflict	conflict with [NAS]
[n] _{/z/}	[NAS]	[COR]	[STR]	[]	conflict	conflict with [STR]

The standard-deviant contrasts are identical in each pair of inversion. The first pair of inversion ([g]_{/d/} ~ [d]_{/g/}) reflects a change in place of articulation, the second pair ([z]_{/n/} ~ [n]_{/z/}) a change in manner of articulation. The shaded area highlights the only completely non-conflicting condition [g]_{/d/} assuming underspecification. MoA = manner of articulation; PoA = place of articulation; [COR] = [CORONAL]; [DOR] = [DORSAL]; [NAS] = [NASAL]; [STR] = [STRIDENT]; [PLO] = [PLOSIVE]; [] = underspecified feature.

Based on the previous studies and following this line of research (Experiment 1; Eulitz & Lahiri, 2004; Walter & Haquard, 2004), the MMN was used again in a classic oddball paradigm in order to rule out additional methodological differences between the experiments of this work. As mentioned in the introduction, the various models of the mental lexicon have different views on how our recognition system copes with this enormous variation in speech, and how detailed phonological information is actually stored in the mental representation. Lexicon models assuming full specification of all kinds of information (Bybee, 2001, Johnson, 1997; Norris, 1994; Norris & McQueen, 2008; Gaskell et al., 1995; Gaskell, 2003) expect to find equal MMNs between all contrasting consonants with varying place features, and manner features since these models assume that all phonological information is fully stored. Remember, that all models but FUL are based on experience - particularly Gaskell and colleagues (Gaskell & Marslen-Wilson, 1996; 1998; 2001) assume that the context is important. However, our stimuli are built in a way that

context does not help – they all have the same vowel contexts. These models all would predict equal MMN activation patterns. Contrarily, models such as FUL (Lahiri & Reetz, 2002), assuming underspecified phonological features predict asymmetries of MMNs within the reversal of consonantal contrasts, presented as standard and deviant when they include [CORONAL] (for summary see Table 7).

With this logic, we expect to find larger MMN amplitudes in conditions with a feature-conflict than in conditions with no conflicting conditions. With respect to the present experiment, we predict a conflict – nonconflict situation for the place contrast, thus expecting asymmetrical MMN responses. This means, for the condition [d]_{/g/} (labeling see Table 7) that the extraction of the underspecified place feature [CORONAL] from the deviant [d] would map onto the specified place feature [DORSAL] in the mental representation created by the standard /g/ and would cause a conflict. In the reversed condition [g]_{/d/} when the feature [DORSAL] from the deviant [g] maps onto the pre-activated representation of the standard /d/ which contains no information about the place of articulation, a non-conflicting situation occurs. Consequently, the MMN amplitude difference between these acoustically equalized conditions [d]_{/g/} (conflict) vs. [g]_{/d/} (non-conflict) differ, with the conflict condition [d]_{/g/} eliciting a larger amplitude, i.e. an asymmetric MMN pattern would again be expected.

Our manner contrast not only helps to test whether the MMN is a sensitive measure for manner-of-articulation features, but also serves as a control, requiring contrasts, which are predicted to be symmetric. Stimuli containing [n] and [z] sounds were used, since both features, [NASAL] and [STRIDENT], are assumed to be fully specified for their manner of articulation in all models under consideration. In doing so, we created a manner conflict in both, the [n]_{/z/} as well as the [z]_{/n/} condition which should result in similar MMN amplitudes. The further question we raise here with both phonological contrasts is whether manner and place conflicts evoke similar brain responses.

3.2.1.1 Methods

3.2.1.1.1 *Participants*

A total of 29 subjects took part in this EEG study. Data from three subjects had to be excluded due to excessive eye movements (EOG), cardiac or muscle artifacts. Thus, the following analysis is based on the remaining 26 participants (age range 20-32 years, mean age 24.92 (SEM .65); 50% female).

3.2.1.1.2 Stimuli

The experimental stimuli used as standard and deviant were four VCV-syllables such as [edi], [egi], [eni] and [ezi] (Figure 13) which are all nonwords in German. Both initial and final vowels are acoustically long (/e:/ and /i:/). Multiple repetitions of these stimuli, along with the disyllable [epi], were recorded by a male speaker. The stimuli were recorded on a digital recorder at a sampling rate of 44 kHz.

Oscillogram

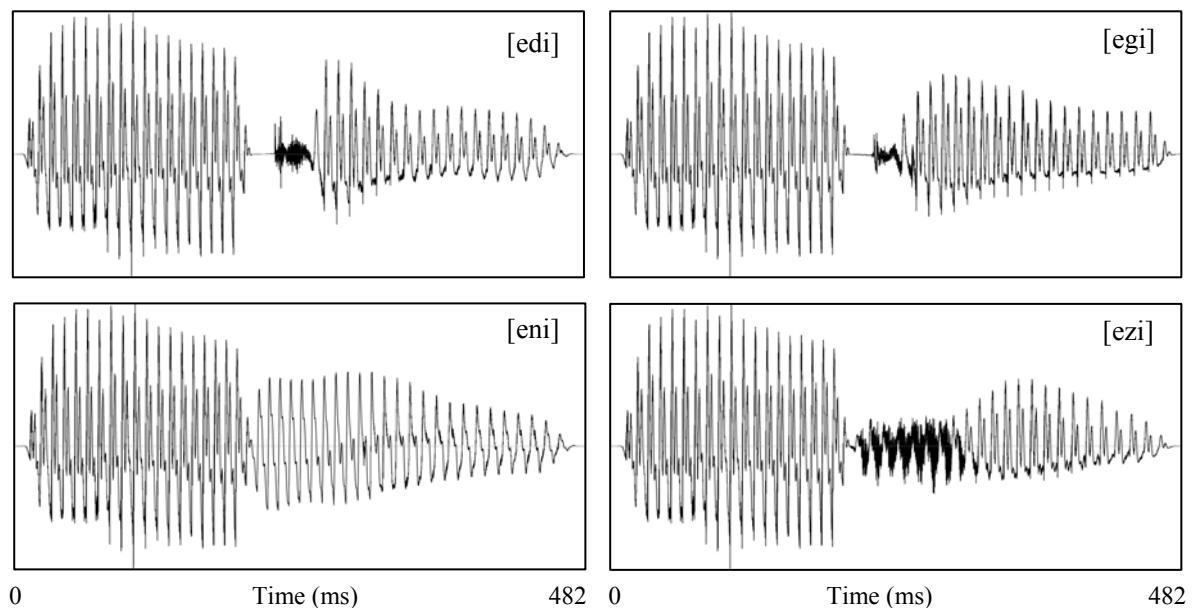


Figure 13: Oscillograms of the acoustic stimuli used in Experiment 3: [edi], [egi], [eni] and [ezi]. The standard and deviant stimuli were maximally matched for their acoustic properties in all conditions.

Three variants of each stimulus type with matched F0 frequencies were selected. Three [e] vowels from the syllable [epi] were cross-spliced on to three variants of the syllables [-di], [-gi], [-ni] and [-zi]. The [e] vowels were taken from the syllable [epi] to prevent further coarticulation effects since [p] has a different place of articulation from the critical consonants. This resulted in nine VCV-syllables, all with the same three [e] vowels attached to three variants of the second syllables in order to reduce acoustic differences between the stimuli. Additionally, pitch and intensity measures were controlled for as best possible. Figure 14 illustrates that each set of VCV-syllables shows comparable variation in pitch and intensity. The plosive sounds are very similar, whereas the nasal and strident sounds differ substantially in pitch and intensity.

All items were controlled for length: 200 ms for [e], 281-283 ms for the second syllable, yielding a total duration of 481-483 ms. It is important to note that the consonantal

onsets vary between the [PLOSIVE] consonants ([d], [g]) and the [NASAL] [n] and [STRIDENT] [z] sounds. The place and manner-of-articulation information of [n] and [z] begin right after the offset of the [e] (after 200 ms) as is common in natural speech. However, the two plosives, [g] and [d] begin with a period of silence after the offset of [e]. Although these are ‘voiced’ sounds, there is no pre-voicing in German. This closure duration lasts 20 ms before the burst onset; consequently, the onset of the place and manner information begins slightly later compared to the [n] and [z] (see Figure 13). However, since the two plosive sounds, as well as the [NASAL] and [STRIDENT] sound are being compared to each other, but not between each other, the comparison remaining within the same category, these latency differences should not create a problem.

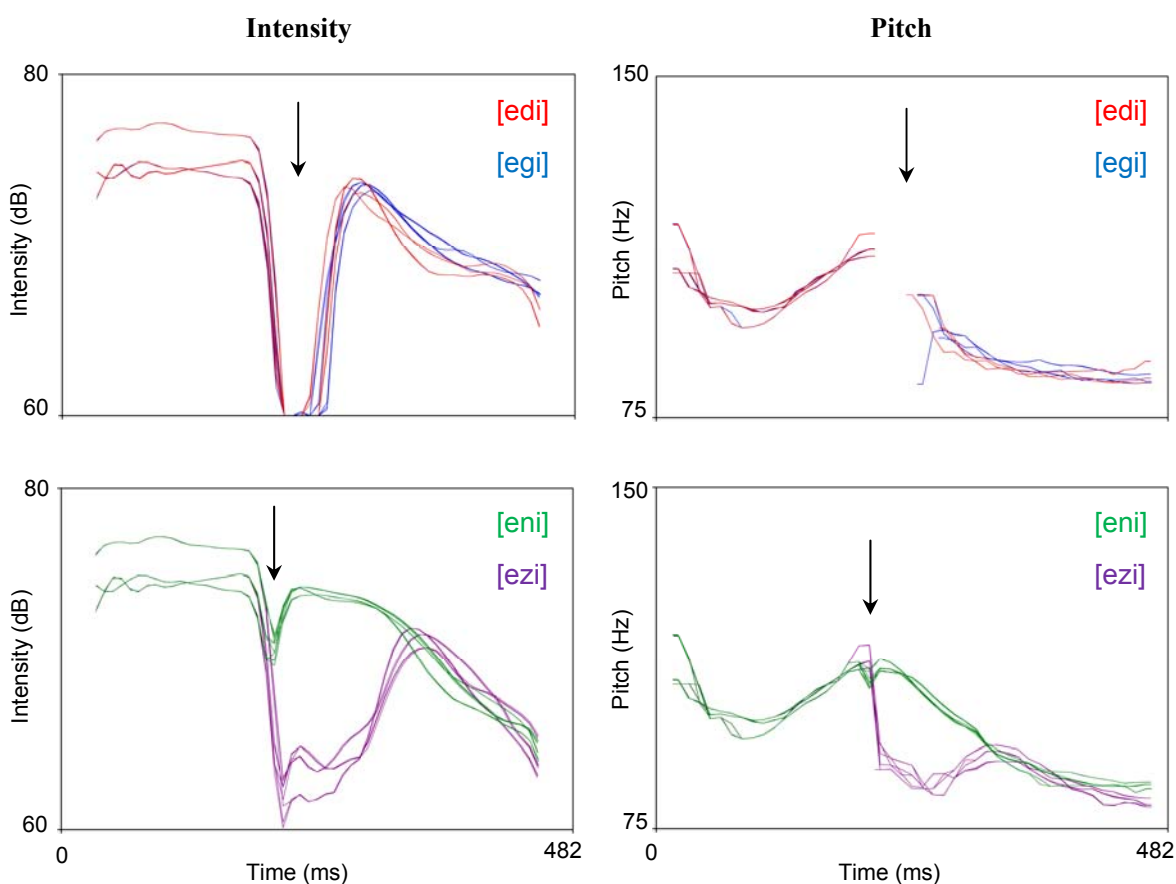


Figure 14: Pitch and intensity contours of the VCV-syllables used in this study: [eni] (green), [ezi] (violet), [edi] (red) and [egi] (blue). Three variants for each VCV-syllable were used, each set of syllables showed comparable variation in pitch and intensity. The plosive place contrast [edi] ~ [egi] is very similar in pitch and intensity, the manner contrast [eni] ~ [ezi] differ quite substantially in pitch and intensity. Note, that the three [e:] sounds are the same across all four VCV-syllables. The arrows mark the onsets of the critical consonants.

All stimuli were equalized for intensity (RMS) to ensure minimal acoustic differences between the stimuli. Since the RMS varied naturally between the experimental

stimuli, an additional loudness rating was conducted to ensure identical perception between the consonants. For this purpose, 27 subjects (none of whom participated in the ERP study) listened to all contrasts via headphones with different manipulated amplitudes (e.g. [eni] with 70 dB was contrasted with [egi] with 66 dB, 68 dB, and 70 dB and vice versa). Subjects were asked to judge whether each pair sounded equally loud. Based on these perception results, the following amplitudes were chosen: 74dB for the initial vowel [e], 68dB for [di] and [gi], 70dB for [ni] and 66dB for [zi]. Additionally, the intensities of the first and last 20 ms of each stimulus were reduced linearly (linear fade in, linear fade out; Adobe Audition) to ensure a smooth beginning and ending.

Again three variants of each stimulus type were used to introduce some acoustic variability to simulate more natural speech perception.

All four sounds were checked for their word-medial frequencies (Table 8). The consonantal sounds differ in their frequency counts in medial intervocalic position, with [n] and [g] being fairly similar ([n] log-value 3.74, [g] log-value 3.26), both being slightly higher than [d] (log-value 2.84), and [z] with the lowest frequency (log-value 0.90) (first column in Table 8). Additionally, distributional frequencies of the speech sounds (phonotactic probabilities), were checked for the first critical vowel-consonant ([VC]) syllables (third column in Table 2). The [VC] sequences were fairly similar, [en] with the highest probability and [ez] with the lowest probability (log-values: [en] 2.81; [ed] 2.25; [eg] 2.10; [ez] 1.89). Even words with initial and final consonants, i.e. in the environment of C[VCV]C, the frequencies of the critical consonants hardly differ from each other (fourth column in Table 8).

Table 8: Log-values of the frequency counts of phonemes and phoneme sequences in the studied VCV-syllables:

nonword stimuli	V[C]V	[VC]V	C[VCV]C
[eni]	3.74	2.81	5.06
[ezi]	0.90	1.89	4.75
[edi]	2.84	2.25	4.73
[egi]	3.26	2.10	4.86

Frequency counts were weighed for word frequency and are based on token counts of the CELEX corpus. [...] indicates the phonemes for which counts are given in the table; [C] = consonant; V = vowel; e.g. C[VCV]C in the fourth column gives the frequency counts for /n/, /z/, /d/ and /g/ in an intervocalic position with further sounds before and after the vowels (e.g. words such as *G[eni]us* ‘genius’ or *G[enu]s* ‘genus’).

3.2.1.1.3 *Experimental Procedure*

In each experimental session the four nonwords were combined in two pairs, with each nonword serving as a standard as well as a deviant, resulting in four blocks (Table 7). During the study, 680 standards (85%) and 120 deviants (15%) per nonword contrast and block were presented.

3.2.1.1.4 *Data Acquisition and Analysis*

For each participant, the averaged MMN responses contained at least 75 % accepted deviant trials in each condition. The MMN was obtained by subtracting the ERP response of the standard from that of the equivalent deviant stimulus across blocks, e.g. [edi] as deviant minus [edi] as standard of the reversed block, resulting in the identity MMN.

Based on visual inspection of the grand-average waveform, the maximum amplitude of the MMN difference waveform of every condition was used for the quantification of the ERPs. The MMN was expected 150-250 ms after the offset of the first vowels (change onset), thus in a time window around 350-450 ms.

For statistical analysis, the individual MMN peak latency and mean amplitude at the Fz electrode are used as dependant variables. The statistical analyses were restricted to paired comparisons and *t*-test because more complex factorial designs were not fully balanced. Full balancing was impossible for the present study where different featural dimensions are varied and other possible confounds (physical differences between stimuli, number of conditions to signal to noise relationship, etc.) had to be avoided. Furthermore, the statistical model was restricted to two independent variables, the two pairs of inversion with an equal acoustic change subjected to two paired *t*-tests (*t*-test, α level = .05) with the pair of inversion PLACE CHANGE ([g]_{/d/} ~ [d]_{/g/}), and MANNER CHANGE ([z]_{/n/} ~ [n]_{/z/}). The *t*-values together with the uncorrected degrees of freedom and the corrected *p*-values are reported.

3.2.1.2 Results

A summary of the averaged waveforms for all four experimental conditions is shown in Figure 15 (see also Appendix for more detailed figures of the grand-average waveforms of standards, deviants and difference waveforms including their maps in Figures 5.1.5a & b). MMN responses in the expected time window (350-450 ms) of the grand average difference waveforms were significant in all four conditions, the largest for the Fz electrode, but with enormous variations in size ([d]_{/g/}: *t* (25) = -5.37; [g]_{/d/}: *t* (25) = -2.13; [n]_{/z/}: *t* (25) = -10.38; [z]_{/n/}: *t* (25) = -10.18).

The MMNs for the asymmetrical contrast PLACE CHANGE ($[g]_{d/} \sim [d]_{g/}$) were significantly different ($t(25) = 2.55; p < .017$). As predicted by the underspecification hypothesis, the asymmetrical pair of inversion for place of articulation showed a larger MMN in the conflict condition, when the place feature [CORONAL] of the deviant [d] mapped onto the pre-activated specified feature [DORSAL] of the standard /g/ in the underlying form. A smaller MMN amplitude in the reversed condition $[g]_{d/}$ was found, when the [DORSAL] deviant mapped onto the standard /d/, which is assumed to be underspecified for [CORONAL]. The MMN amplitudes of the bidirectional conflicting manner conditions MANNER CHANGE ($[n]_{z/} \sim [z]_{n/}$) did not differ from each other ($t(25) = -0.43; p > .1; n.s.$) as expected, having equal manner conflicts in both directions of presentation, one conflict for the feature [NASAL] and one for [STRIDENT].

The interesting point is whether the above reported differential MMN effect between the PLACE CHANGE and the MANNER CHANGE is systematic in nature. To examine this question, the differences between the contrasts were calculated and statistically compared. The analyses revealed a significant difference between PLACE CHANGE and the MANNER CHANGE contrast ($t(25) = 2.35; p < .026$), showing largely enhanced MMN amplitudes for the MANNER CHANGE contrast (Figure 15 & 16; see Table 16a of 5.2.2 Statistical Tables in the Appendices for detailed statistical analysis).

Across all experiments the same amplitude analysis was calculated for condition-related mean MMN amplitudes for Fz and Cz to ensure consistency. This condition-related analysis was measured separately for each grand-averaged condition, resulting in four grand-average time windows for each condition. The analysis showed similar results as for the individual mean MMN amplitudes but less pronounced (see Appendices 5.2.2 Statistical Table: Table 16a for statistical analysis).

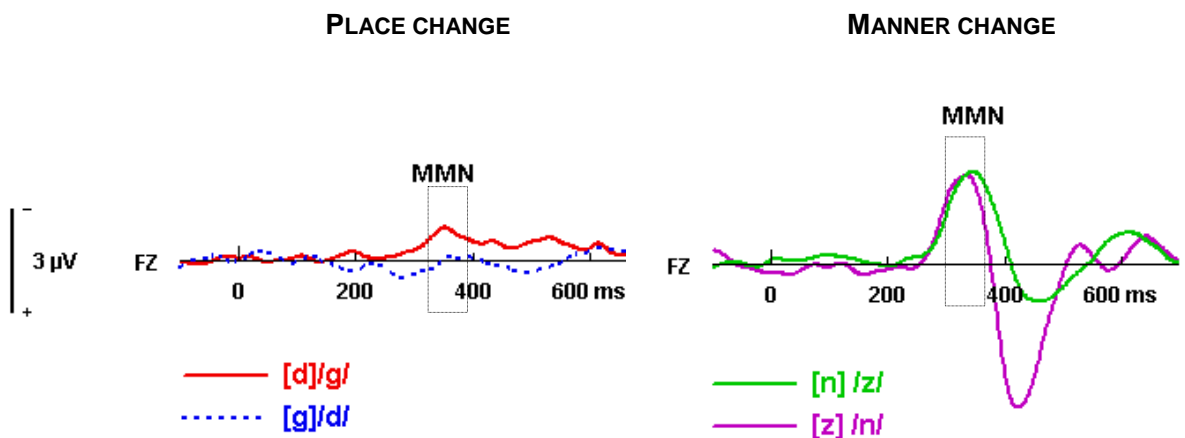


Figure 15: Grand-average difference waveforms of all four conditions (re-referenced against the right mastoid (Rm), the PLACE CHANGE pair of inversion ($[d]_{g/}$ (red) & $[g]_{d/}$ (blue, dashed)), and the MANNER CHANGE pair ($[z]_{n/}$ (violet) & $[n]_{z/}$ (green)) at the Fz electrode. Note, that the only non-conflicting condition is shown in dashed lines. The square areas mark the time window of the MMN.

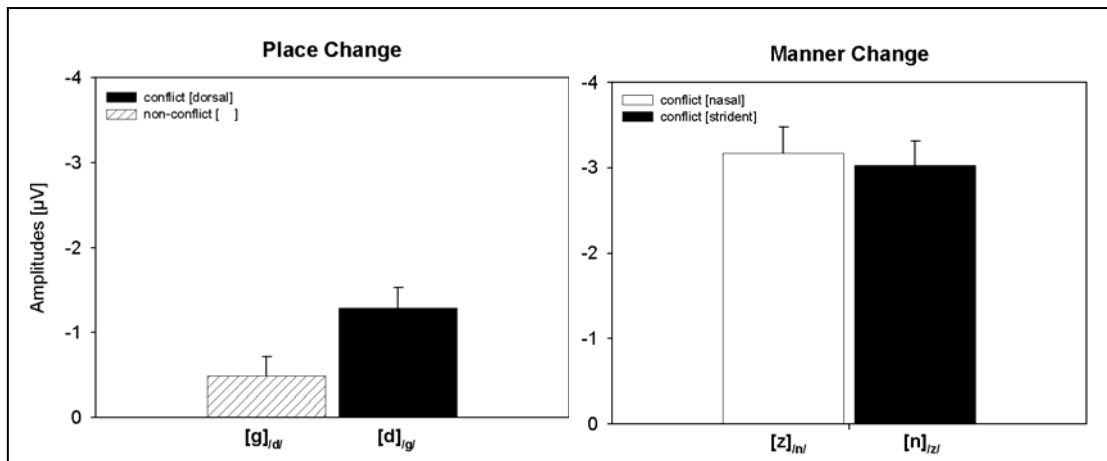


Figure 16: MMN mean peak amplitudes of all four conditions at Fz, the PLACE CHANGE and the MANNER CHANGE contrast. The striped bar indicated the only non-conflicting condition. Error markings on top of the bars represent the standard error of the mean (SEM).

Latencies were further analyzed for the PLACE CHANGE and MANNER CHANGE contrasts. The MMN latencies of the two plosive sound conditions did not differ from each other as expected ($t(25) = 1.58; p > .1; n.s.$). Similarly, there was no significant difference for the MANNER CHANGE contrast although the [STRIDENT] sound in the function as deviant tended to be slightly earlier ($t(25) = -1.93; p < .065; n.s.$; see Figure 15 (Manner change) and Appendices 5.2.2 Statistical Table 16b for statistical analysis).

P300 Effect

Another unexpected effect surfaced in the data the so-called P300 effect. Based on visual inspection of the grand-average waveforms, large positive deflections at fronto-central electrode sites were observed immediately following the MMN for the two manner-of-articulation conditions (see Figure 15 (Manner change); also Appendices Figure 5.1.5b). This P300-like component was largest at Fz and Cz and rather small at Pz, which could reflect a P3a having more frontal distribution (Squires et al., 1975; Knight et al., 1989; Katayama & Polich, 1998). Therefore, further analyses were carried out. Due to its frontal location it was not clear whether this positivity indeed reflected a subcomponent of the P300, such as the P3a or just a repolarization phenomenon of the MMN. For this reason this positivity was parameterized in the same way as was done for the MMN, which will be described in the following:

Amplitude measurements of the P300 are usually taken at one electrode, typically at either Cz, or Pz (Picton, 1992); however we looked at the Fz (as done for the MMN measurements). For statistical analysis, the mean amplitude of the Fz electrode was used as a dependant variable which was measured over a 50 ms time window from positivity onset to

offset (400-450 ms) and thereby well covering the largest positive peak and its surrounding area.

The statistical analyses were restricted to paired comparisons and *t*-test with the two pairs of inversion PLACE CHANGE ($[g]_{/d/} \sim [d]_{/g/}$), and MANNER CHANGE ($[z]_{/n/} \sim [n]_{/z/}$). Amplitude values of the difference waves were also tested against zero in two-tailed *t*-tests separately for each condition.

Only the two manner conditions revealed significant positivities ($[n]_{/z/}$: $t(25) = 2.21$; $p < .037$; $[z]_{/n/}$: $t(25) = 8.45$; $p < .000$), the two place conditions did not show any positivity in the typical P300 time window ($[g]_{/d/}$: $t(25) = 0.25$; $p > .1$; n.s.; $[d]_{/g/}$ revealed a significant negativity: $t(25) = -2.21$; $p < .036$). The positivities for the two manner conditions MANNER CHANGE ($[n]_{/z/} \sim [z]_{/n/}$) differed significantly from each other ($t(25) = 6.24$; $p < .000$). The P3a-like component was much larger for the condition $[z]_{/n/}$, compared to $[n]_{/z/}$ (see Appendices 5.2.2 Statistical Table 16c for detailed statistical analysis).

3.2.1.3 Discussion of Experiment 3

The present MMN study examined the fine structure of phonological specifications with respect to the underspecification of the feature [CORONAL] with stop consonants, and the sensitivity of the MMN using manner features. This was done to make a further step in generalizing the results for place of articulation in vowels to other sounds such as stops (in medial word position) and feature dimensions (manner of articulation). The underspecification of the [CORONAL] place of articulation has been demonstrated in other MMN studies for isolated vowels (Eulitz & Lahiri, 2004) and vowels embedded in linguistically more complex structures (Experiment 1), as well as for sounds in different word positions (e.g. medial word position: Friedrich et al., 2006; initial word position: Friedrich et al., 2008; Walter & Haquard, 2004). When embedded into vowel contexts, the consonantal contrasts under study showed supporting evidence for the underspecification of the [CORONAL] place of articulation in stop consonants. Larger MMN amplitudes were found when a place-of-articulation conflict occurred compared to a non-conflicting condition. According to FUL, a conflict in this study occurs (Lahiri & Reetz, 2002) when the [CORONAL] place-of-articulation information is extracted from the acoustic signal by the deviant [d] among a row of standard stimuli, i.e. /g/, which have pre-activated the [DORSAL] place of articulation. In the reversed case, when [g] is the deviant and /d/ the standard no conflict is expected. Models that assume fully specified phonological information or models that depend on experience and contextual information (Bybee, 2001; Gaskell et al., 1995; Gaskell, 2003; Gaskell & Marslen-Wilson, 1997; 1996; 1998; Norris, 1994; Norris & McQueen, 2008) would predict similar conflicts for all directions of change resulting in

similar MMN response patterns. Our present findings are in line with the predictions of the FUL model. We found the expected asymmetrical MMN pattern for the place-of-articulation conflict ~ non-conflict contrast ($[d]_{/g/} \sim [g]_{/d/}$), a higher amplitude when the coronal stop maps onto the dorsal stop in the underlying representation, and symmetrical MMN amplitudes for our manner condition, the conflict ~ conflict contrast ($[n]_{/z/} \sim [z]_{/n/}$). Interestingly, the phonetic contrasts evoked MMNs, which differed in the two feature dimensions under investigation. The manner contrast produced much larger MMNs than the place contrast, which also shows that the MMN indeed is a sensitive measure for studying manner features of articulation. This might have been caused by differently located (or oriented) neural generators. Another possible explanation for the significantly smaller MMNs of the place contrast might be due to the fact that they are stop consonants, which are rather weak sounds. Previous studies examining phonetic contrasts between stop consonants manipulating mainly place of articulation (Maiste et al., 2005; Shafer et al., 2004; Diesch & Luce, 1997; Pettigrew et al., 2004a, 2004b) back up this finding. For example, Pettigrew et al., (2004a) found poor MMN responses to their fine acoustic speech contrasts $[de/ge]$ and $[day/gay]$ proposed that when carefully controlled methodological designs and strict methods of analysis are applied, robust MMN responses to fine-grained phonetic contrasts may be difficult to obtain. Studies by Dalebout and Stack (1999) and Dalebout and Fox (2000) also found that MMN responses to fine acoustic contrasts such as $[d/g]$ may not always be obtained. Nevertheless, we do find small MMNs for our fine-grained place feature contrasts in the medial position with the predicted asymmetrical response pattern, which suggests that even in minimal acoustic contrasts the assumed difference in mental representation can be detected. This however was not found in Experiment 2.

The MMN has been shown to be sensitive to many different factors such as pitch changes, as well as intensity, duration, and inter-stimulus interval differences (for review see Näätänen, 2001). Our study controlled for all these factors not only within the stimuli but also by using the identity MMN approach (Eulitz & Lahiri, 2004; Pulvermüller & Styrov, 2006). Is it still possible to account for these results purely based on acoustic characteristics of these stimuli? The plosive sounds $[d]$ and $[g]$ in our place contrast are quite similar acoustically (pitch and intensities, see Figure 14). This is different for the nasal $[n]$ and the strident $[z]$ in our manner contrast, which differ acoustically to a great extent. This would lead us to the prediction, that the greater the acoustic difference between the stimuli the bigger the difference in MMN amplitudes will be. However, our results do not show this pattern: Our acoustically similar pair of plosives elicited different MMN amplitudes, with the deviant $[d]$ showing a higher amplitude than the deviant $[g]$, while our acoustically dissimilar pair elicited equal MMN amplitudes. These findings support our claim of asymmetrical representations that $[NASAL]$ and $[STRIDENT]$ are both specified in the mental representation

and that these features conflict with each other, whereas [CORONAL] and [DORSAL] are not represented in the same way.

Among other possible parameters, which could have affected our results, the influence of phonotactic probabilities could be an issue. Bonte et al. (2005) reported possible effects of phonotactic probabilities on the MMN. They found higher MMN amplitudes for high probability nonwords as compared to low probability nonwords. The distributional probabilities of our [VC] sequences in this experiment were lowest for the cluster with [z], all other stimuli were fairly similar (see methods and Table 8). Comparing Bonte et al.'s findings to our study, higher and fairly equal MMN amplitudes would be expected for [edi] and [egi], as well as [eni], and a lower amplitude for the [ezi] condition. Yet this is not what we found. The conditions with the fairly similar phonotactic probabilities showed MMN differences ([edi] > [egi]) whereas the conditions with the largest difference in phonotactic probabilities showed no MMN difference ([ezi] = [eni]). So, contrary to Bonte et al. (2005), no differences in MMN amplitude were found between our high vs. low phonotactic probability condition, yet a difference was found for our two conditions with similar phonotactic probabilities.

Furthermore, our findings cannot be explained by individual sound frequency effects of our stimuli. The intervocalic frequency turns out to be highest for [n] and [g], slightly lower for [d] and lowest for [z] (Table 8). Again, one could argue that a higher frequency causes a higher MMN response; however in our results the MMN amplitude of the deviant [d] with the lower frequency value compared to the standard /g/ is increased compared to the reversed condition. Additionally, the largest sound frequency difference is between [n] and [z], but here we find equal MMN amplitudes. These results show a pattern, which cannot be explained by frequency effects, or phonotactic probability influences.

P300 – Automatic Attention Switch

The results also revealed the presence of a robust positive component immediately following the MMN for the two manner-of-articulation conditions. This effect was more evident at the Fz and Cz than at parietal regions. Therefore this positivity was identified as P3a component due to its frontocentral distribution and by being elicited without a specific task assignment given to the subject. The P3a frequently follows the MMN (Lyytinen et al., 1992) and is proposed to reflect the neural correlate of an orienting response (Soltani & Knight, 2000; Friedman et al., 2001), an automatic attention switch toward unattended changes in the stimulus context when the primary task is not very demanding (Sams et al., 1985a; Escera et al., 2000). The P3a component is thought to reflect the activation of a cerebral network in involuntary orienting of attention toward unattended changes in the stimulus context (Escera et al., 2000). A larger P3a amplitude was elicited by the condition where the standard was a

[NASAL] /n/ sound and the deviant was a [STRIDENT] [z]. This manipulation of deviant / standard sound stimulus context produced a task environment in which the deviant sound [z] might to have engaged focal attention similar to the one observed for highly ‘novel’ stimuli. The frontal maximum deviant P300 of the present study can, therefore, be considered as a manifestation of the P3a, which may be the same as the novelty P300 (Squires et al., 1975). The neural representation of the stimulus environment provides the ‘context’ upon which attentional processes operate. It can be speculated that if this context is defined primarily by a non-sonorant deviant and sonorant standard discrimination as in the case of the [z]_{/n/}, attentional redirection to the quite different [STRIDENT] sound occurs. This perhaps demands more attentional effort activating more frontal lobe activation that may underlie the P3a component than the reversed case where the [NASAL] deviant stimulus could more easily be ignored. The non-sonorant [STRIDENT] sound seems to be more salient (as ‘novel’ sounds) suggesting, that when the acoustic qualities of sounds differ radically from the qualities of the background stimuli (standards) – in our study perhaps the difference between [NASAL] and [STRIDENT] – the P300 effect reflects the orienting response in the first place with no distinct differentiation of the sound nature per se. This could be true, since the manner change, this [NASAL] ~ [STRIDENT] contrast might differ more than the place change thus requiring more attentional resources. This could have led to the P3a deflection for the manner contrast ([NASAL] ~ [STRIDENT]) and not for the place contrast ([CORONAL] ~ [DORSAL]). According to Squires et al. (1975), the P3a reflects a ‘mismatch’ to an ongoing stimulus train. In this interpretation, the P3a would index a basic sensory mechanism, which registers any change in a background stimulus, perhaps by means of mismatching a specific neural ‘model’ (Sokolov, 1963) established by repetition of the background. In the case of our data both manner conditions seem to mismatch stronger, registering a change more strongly compared to our place change where no P3a was elicited. It was even stronger in the case where the deviant is a [STRIDENT] compared to a [NASAL]. It seems that the [STRIDENT] attracts more attentional resources.

In conclusion, MMN asymmetries were found for pairs of consonants with varying place features such as [DORSAL] ~ [CORONAL] taking a further step in generalizing the underspecification of the feature [CORONAL] to other speech sounds. Furthermore, MMN symmetries were found for pairs of consonants with varying manner features ([NASAL] ~ [STRIDENT]) examining manner features in a controlled way in an MMN study for the first time. It has been shown that the MMN seems to be a reliable and sensitive measure to study these other feature dimensions, which in effect is the basis for further research in this direction. The place contrast showed the expected asymmetry which supports phonologically underspecified mental representations during speech perception.

3.2.1.4 General Discussion of Place Features for Stops and Manner Features

Our experiments were aimed at exploring how the brain encodes speech sounds differing in phonological features, and how these different features might be represented. More specifically, the two previous Experiments 2 and 3 tried to examine phonological specifications with respect to the place-of-articulation features [LABIAL], [CORONAL], and [DORSAL] in stop consonants, as well as manner-of-articulation features [NASAL] and [STRIDENT]. Both studies varied stop consonants in different positions contrasting place of articulation, in an initial position in Experiment 2 and in a medial position in Experiment 3. Additionally the third study also contrasted the manner of articulation using nasal and strident consonants. A main question was whether the MMN indeed is sensitive enough to test manner-of-articulation features and their mental representations.

Experiment 3 resulted in asymmetrical MMN activation patterns for the place-of-articulation contrast [CORONAL] ~ [DORSAL] for stops in word medial position. A higher amplitude was found when the coronal stop was the deviant and mapped onto the dorsal stop as standard, conflicting with the underlying representation compared to the reversed non-conflicting condition where a smaller amplitude was found. This pattern supports phonologically underspecified mental representations during speech perception. The manner-of-articulation contrast was assumed to conflict in both directions since [STRIDENT] and [NASAL] are both specified and it was assumed to result in symmetric MMN responses. The MMNs for this manner contrast showed exactly this pattern. Additionally, this showed that the MMN can be used as a sensitive measure for examining manner-of-articulation features, which in effect is the basis for further research in this direction. In Experiment 2, the place contrasts, with stops in syllable initial position produced significantly smaller MMN responses. None of the expected asymmetries were significant.

Both our experiments have found very small or no MMN responses to place-of-articulation contrasts with stops which is in line with previous literature (e.g. Shafer et al., 2004; Diesch & Luce, 1997; Pettigrew et al., 2004a; 2004b; Wunderlich & Cone-Wesson, 2001; Dalebout & Stack, 1999; Dalebout & Fox, 2000). The fact that the stop contrasts in both experiments produced such small MMN responses might be due to them being very short, context-dependant and weak in themselves. They might indeed be more sensitive to interference by noise, such as acoustic or internal (Cowder, 1971), that is, being less discriminable than vowel contrasts and other non-plosive consonants. It is suggested that the auditory memory traces that are created by stops are weaker than the traces created by vowels. This consonant-vowel difference was obtained both with syllable-initial and syllable-final consonants (Cole, 1973) and might affect the amplitude of the mismatch responses, as it might be the case especially for our second experiment. The MMNs of Experiment 3 were elicited by stops in medial position, showing actually larger MMNs than

those in Experiment 2 with stops in syllable-initial position. The stops of Experiment 3 were embedded into vowel contexts automatically leading to more contextual information, which might be necessary for the perceptual system to detect the fine phonological feature differences. Language in itself is contextual, thus when using speech sounds without much contextual information it might be difficult for the perceptual system to detect the differences. The stimuli in Experiment 2, despite being natural speech as opposed to synthesized speech, might have been too abstract and too short as speech input. These results are in line with studies by Shafer et al. (2004) and Maiste et al. (1995) observing MMNs less than 1 μ V to their stop consonant contrasts, arguing for difficulty in discrimination that might have led to a small magnitude MMN which cannot be seen without greater signal/noise ratio. Similarly, we can conclude that our stop contrasts especially in Experiment 2 were more difficult to process than other sounds. Despite the fact that these studies elicited small MMN responses, they remain interesting since they nevertheless show sensitivity towards phonological contrasts (Dehaene-Lambertz, 1997; Rivera-Gaxiola et al., 2000; Shafer et al., 2004; Maiste et al. 1995). For instance, Shafer et al. (2004) looked at the categorical perception of place of articulation in plosives of English and Hindi speakers, using an oddball paradigm. A stronger MMN was found for their labial-dental contrast for both language groups when the [LABIAL] sound /b/ served as a standard and the [DENTAL] sound [d] functioned as a deviant, but not reversed. This asymmetry of MMNs supports our assumptions of the underspecification hypothesis. A further MMN study, which supports the assumptions of underspecification, was conducted by Maiste et al. (1995). She examined the effects of categorical perception on place of articulation using voiced stops in CV-syllables, which were computer-modified speech sounds from a nine-stimulus continuum between /ba/ and /da/, changing mainly the formant transition in the second formant (F2). She manipulated within category contrasts ([LABIAL] ~ [LABIAL] and [DENTAL] ~ [DENTAL]) and across category contrasts ([LABIAL] ~ [DENTAL]) in a passive oddball paradigm. Her second, third, and fourth experiments are interesting in terms of our underspecification assumption. In her second experiment, the across-category contrast ([da]_{/ba/}) produced a larger MMN than the reversed across-category contrast ([ba]_{/da/}); the within-category contrasts ([ba]_{/ba/} ~ [da]_{/da/}) either resulted in no MMN or very small MMNs. The third experiment used the prototypical endpoints of the stimulus continuum as standards and all others as deviants. Acoustically it was found that the farther away the deviant was from the standard, the larger the MMN was. Additionally, the conditions with /ba/ as standard resulted in larger MMNs than with /da/ as standard. The MMN for the across-category contrast [da]_{/ba/} produced a much larger amplitude compared to the reversed condition. The fourth experiment again revealed the same pattern within the phonetic MMN, a larger MMN for the /ba/ standards than for the /da/ standards. Maiste et al. (2004) tried to explain her results in light of

frequency differences of the stimuli, and reduced her MMN asymmetries purely to the detection of acoustic changes. These results by Maiste and her colleagues, together with the results of Sharfer et al. showed larger MMN amplitudes for the phonological conflict conditions ([d]_{/b/}) as compared to those conditions with no conflict ([b]_{/d/}). Their results however are not discussed in light of asymmetrical representations as we propose. Their asymmetries of MMNs can be very well explained by the underspecification of the [CORONAL] place of articulation.

These studies all show that an MMN can be observed when contrasting place of articulation with stop consonants, and thus the MMN is sensitive to explore the phonological status of stops. Nevertheless, these studies also show very heterogeneous results when using place contrasts with stops as partly described above (Deheane-Lambertz, 1997; Rivera-Gaxiola et al., 2000; Shafer et al., 2004; Maiste et al., 1995; Aulanko et al., 1993; Diesch & Luce, 1997; Pettigrew et al., 2004a, 2004b; Dalebout & Stack, 1999; Dalebout & Fox, 2000; Sharma et al., 1993; Wunderlich & Cone-Wesson, 2001).

One major methodological difference of all these studies and our present two studies with stops, as well as the study by Eulitz and Lahiri (2004) and Experiment 1, is the approach of how the MMN is analyzed. In our studies, the iMMN was used, which is based on a comparison of identical stimuli across blocks, in their roles as standard and deviant. The other studies used a direct comparison between the standard and deviant within the same block, thus comparing different stimuli with each other. A direct comparison of iMMN results and the classic MMN results remains problematic.

A further reason that may have led to the diverging results of Experiments 2 and 3 might lie in the method of the MMN paradigm, causing less reliable results for Experiment 2. A passive roving-standard paradigm was used here, whereas a classical constant-standard oddball paradigm was used in Experiment 3. In the roving-standard design, features of the standards (and deviants) change between stimulus trains, in that less useful information could be assumed to have developed over the session not having activated the necessary underlying representation.

Winkler and his colleagues, (1996) tested predictions of a memory-representation hypothesis of the MMN by comparing the constant-standard paradigm and the roving-standard paradigm. In their roving-standard condition, they presented trains of tones, with the standard frequency changing from train to train. In the constant-standard condition, they presented both standard and deviant as fixed sequences throughout the session. Besides the finding that the MMN was elicited after two or more standards, that is, by the deviant in third position or later in the roving-standard condition, they also found that within the constant-standard condition the MMN amplitude increased from the beginning to the end of the stimulus blocks. They concluded that this might be due to the cumulative effect of the

repetitions of the constant standard stimulus (mediated by its memory trace), which is not seen in the roving-standard conditions. This cumulative effect of the repetitions could have created a sharper memory representation for that stimulus, thus enhancing the MMN (Winkler et al., 1996).

The fact that we used the roving-standard paradigm in our second experiment might have influenced the results in exactly this way, not building this cumulative effect of repetitions of the standard and not creating strong and sharp enough memory representations in order to elicit reliable MMNs and to show representational differences between the sounds. Additionally, the stimuli in these studies were tones and not language. It might be hypothesized that in order to build a strong enough representation through the standard for language sounds, more standards indeed are necessary in order to tap into lexical representations and to activate differences in their representations

In conclusion, this third experiment provides electrophysiological evidence for the notion of underspecified and specified cortical representations for place and manner of articulation in medial position for nonwords, finding an asymmetry between [CORONAL] and [DORSAL] speech sounds and a symmetric activation pattern between [NASAL] and [STRIDENT]. Our study with stop consonants in initial position (Experiment 2) did not show any differences and therefore did not provide any further information about mental representations. These results might have been caused by methodological issues or by the stimulus material itself as discussed above. The fact that the results on plosive stops seem so heterogeneous remains an open issue to date.

Our results nevertheless cover new grounds by contrasting place of articulation with stop consonants and manner of articulation, generalizing the previous findings of asymmetrical activation patterns to other feature dimensions. If the sounds were intrinsic wholes, as assumed by many other models (Bybee, 2001; Johnson, 1997; Norris, 1994), we should not expect any differences or only differences based on frequency or contextual effects, unless there really are differences in their underlying representations.

In looking at other feature dimensions such as manner of articulation, as well as place of articulation with stops in medial position, we are able to provide strong support for the assumption that the brain refers to more abstract underspecified phonological representations and does not reflect a one-to-one relation between the speech signal and mental representations of sounds.

3.2.2 Experiment 4: Underspecification of the Manner Feature [PLOSIVE]²³

The notion of underspecification is not only relevant for [CORONAL], but is applicable to other features as well (cf. Felder et al., 2008; Scharinger et al., 2009; Scharinger & Lahiri, 2010; for vowel height features). A basic assumption in phonological representation is that neither consonants nor vowels are indivisible wholes. Rather they are made up of featural properties. It is assumed that consonants differ in terms of place and manner of articulation. Not all of these properties have the same status in that asymmetric relationships exist between them. The features [PLOSIVE] and [NASAL] are also asymmetric in their pattern of alternations and it is assumed that this is so in the mental representation. This is first of all assumed, because word medial [PLOSIVE + NASAL] sequences hardly exist in English and German. There are no [bm] sequences in either language, no [dn] sequences in German and very few [gn] sequences in both: *kidney*, *magnet* (Engl.); *Magnet* (German). Second, [PLOSIVE + NASAL] sequences like [nd], [mb] or [ng] can often become all nasal, but not the reverse. One often hears this in American English dialects in words such as, *kindergarten* and *tendency*. Words like *finger*, change in place and nasality: [ng] > [ŋg] > [ŋŋ]. Historically, we find evidence for [PLOSIVE +NASAL] sequences becoming [NASAL], but never [PLOSIVES]. For instance, *number* in English corresponds to *Nummer* in German where the [mb] sequence is totally assimilated; *lamb* in both languages is pronounced without a final [b]. Therefore, it is proposed that [PLOSIVE] is underspecified while [NASAL] is specified.

Neurobiological studies have mainly shed light on the underspecification of the ARTICULATORS, in particular PLACE (Friedrich et al., 2006; Friedrich et al., 2008; Eulitz & Lahiri, 2004; Walter & Haquard, 2004; see also Experiment 1 & 3), but have hardly considered any other featural dimension. The present study is designed to fill this gap and produce evidence for the generalization of the underspecification assumption. The manner features we will focus on here are [NASAL], [STRIDENT], and [PLOSIVE], in addition to the place features [CORONAL] and [DORSAL]. Based on asymmetries in historical and synchronic alternations, we also assume asymmetries in representation and processing. Thus, we assume that [PLOSIVE] is underspecified, while [NASAL] and [STRIDENT] are not.

Former studies have demonstrated that the MMN is a robust measure to study aspects of feature specification of segments (Eulitz & Lahiri, 2004; Walter & Haquard, 2004; Experiment 1 & 3), which is why the MMN was again used in this fourth experiment.

²³ This study has been submitted to the Journal of Experimental Psychology: Human Perception and Performance as Cornell, A.S., Eulitz, C. & Lahiri, A., (submitted). 'Inequality In Manner-Of-Articulation Contrasts In Speech Perception: Evidence From Event-Related Brain Potentials'.

The MMN experiment reported here has several new aspects: (i) class of speech sound contrasts, (ii) featural dimensions, (iii) medial position. In our earlier studies, we examined fine-grained contrasts in vowels. Crucially, certain manners of articulation differences can only be studied using consonantal speech sounds. Although consonants have been used in a variety of MMN experiments (for review see Näätänen, 2001; Pulvermüller & Shtyrov, 2006a), to answer phonologically motivated research questions, fine-grained differences in consonants have only been used in a handful of studies. These MMN studies (Bonte et al., 2005; Maiste et al., 1995; Shafer et al., 2004; Sharma et al., 1993; Walter & Haquard, 2004) however, focused on place-of-articulation differences. In the present study, consonants were manipulated contrasting manner and place that are embedded in identical nonword VCV-structures.

Here we used the sequences [edi], [eni], [ezi] and [egi]. While [edi], [eni], and [ezi] contrast in manner of articulation keeping the place information constant (all stimuli are coronal consonants). Furthermore, [egi] shares the same manner features with [edi] but its place of articulation is [DORSAL]. Since the contrast we were interested in is medial, where no particular assimilatory context could be responsible for the consonantal variation, our stimuli were necessarily disyllabic (these stimuli are the same used in Experiment 3 – however using different contrast configurations).

According to FUL, [PLOSIVE] manner of articulation is underspecified, while the others are specified in the mental representation. The featural details of the phonological representations of the consonants under study are given in column 2 of Table 9. The third column lists the features that are extracted from the signal and the final column lists the predicted levels of conflict between phonological features in surface and underlying representations. For example, [NASAL] and [STRIDENT] conflict with each other (condition (b)). Since [PLOSIVE] is unspecified, when [NASAL] or [STRIDENT] are extracted, they do not conflict with it. This is however, not the case in the opposite directions — [PLOSIVE] conflicts with [STRIDENT] and [NASAL] (condition (a)). Here we predict an asymmetry of MMNs within the reversal of consonantal contrasts, presented as standard and deviant.

In line with other studies (Eulitz & Lahiri, 2004, Experiments 1 & 3), we expect to ascertain larger MMN amplitudes in conditions with a feature-conflict as compared to conditions with no featural conflict. Within the comparisons in this experiment, in the [d]_{/n/} condition (labeling see Table 9) the feature [PLOSIVE], extracted from the deviant [d], would conflict with the lexically specified [NASAL] created by the standard /n/. In the reversed condition [n]_{/d/}, when the feature [NASAL] is extracted from the deviant [n] and mapped onto the pre-activated representation of the standard /d/, which contains no information about the manner of articulation, a non-conflicting situation occurs. Consequently, the MMN amplitude for these acoustically equalized conditions [d]_{/n/} (conflict) and [n]_{/d/} (non-conflict)

should be larger for the conflicting condition compared to the non-conflicting condition, i.e. an asymmetric MMN pattern would be expected (contrast (a)). More symmetry in MMN responses is expected for [n] and [z], since both features, [NASAL] and [STRIDENT] are assumed to be fully specified for their manner of articulation. Consequently, they should conflict in both directions (contrast (b)). Thus, the [n]_{/z/} ~ [z]_{/n/} contrast should result in similar MMN amplitudes which also has already been shown in Experiment 3. Note, that in all conditions mentioned so far, the place of articulation is held constant since they are all [CORONAL] consonants.

Table 9: Predictions about the amount of feature conflict for the nonword stimuli used in the six experimental conditions

	Experimental conditions [Deviant] _{/Standard/}	Features in the surface representation (extracted by the deviant)		Features in the lexical representation (activated by the standard)		Full specification hypotheses	Underspecification hypotheses
		MoA	PoA	MoA	PoA		
a)	[n] _{/d/}	[NAS]	[COR]	[]	[]	conflict	non-conflict with []
	[d] _{/n/}	[PLO]	[COR]	[NAS]	[]	conflict	conflict with [NAS]
b)	[n] _{/z/}	[NAS]	[COR]	[STR]	[]	conflict	conflict with [STR]
	[z] _{/n/}	[STR]	[COR]	[NAS]	[]	conflict	conflict with [NAS]
c)	[n] _{/g/}	[NAS]	[COR]	[]	[DOR]	conflict	conflict with [DOR]
	[g] _{/n/}	[PLO]	[DOR]	[NAS]	[]	conflict	conflict with [NAS]

The standard-deviant contrasts are identical in each pair of conditions. Note, that the [CORONAL] place of articulation is the same for all sounds, but [g] having a [DORSAL] place. The shaded area highlights the only completely non-conflicting condition [n]_{/d/} assuming underspecification. MoA = manner of articulation; PoA = place of articulation; [COR] = [CORONAL]; [DOR] = [DORSAL]; [NAS] = [NASAL]; [STR] = [STRIDENT]; [PLO] = [PLOSIVE]; [] = underspecified feature.

The third condition (condition (c)) explores the MMN activation pattern varying both manner and place of articulation at the same time, contrasting the phonemes /n/ and /g/. The nasal /n/ is only specified for the manner feature and its place feature remains unspecified. In contrast, /g/ is specified for its [DORSAL] place, but not for its manner of articulation. Thus, when the manner feature [NASAL] is extracted from the deviant [n], and mapped onto the underspecified manner feature [PLOSIVE] of the standard /g/, no conflict occurs for manner of articulation. However, the [CORONAL] feature of the deviant [n]

conflicts with the specified place feature [DORSAL] of the standard /g/, creating a PLACE conflict for the condition [n]_{/g/}. For the reversed condition [g]_{/n/}, when the place feature [DORSAL] is extracted from the deviant [g], it does not conflict with the unspecified [CORONAL] place of the standard /n/, but its manner feature [PLOSIVE] conflicts with the specified [NASAL], having a MANNER conflict for this condition. Thus, conflicts are expected in both directions of acoustic change but for different feature dimensions. The question we ask here is whether manner and place conflicts evoke similar brain responses, or if one sort of conflict is likely to lead to larger MMNs. As of yet this is uncharted territory.

Mental lexicon models assuming full specification (Bybee, 2001; Gaskell et al., 1995; Gaskell, 2003; Norris, 1994; Norris & McQueen, 2008) expect to find equal MMNs between all contrasting consonants with varying manner features ([PLOSIVE], [NASAL] or [STRIDENT]) since it is assumed that all phonological information is fully stored. As we mentioned earlier, models other than FUL are based on experience; either all variation is stored (Bybee, 2001; Johnson, 1997), or only forms in their canonical forms are stored (Gaskell & Marslen-Wilson, 1997; Gaskell et al., 1995; Norris, 1994; Norris & McQueen, 2008) and contextually sensitive information compensates for variation (Gaskell and Marslen-Wilson, 1996, 1998, 2001). Since our stimuli are designed in a way that context remains constant – they all have the same vowel contexts – these models would presumably predict equal MMN activation patterns. If anything, differences in MMN would then have to be attributed to acoustic or frequency characteristics of [PLOSIVE], [NASAL] and [STRIDENT].

3.2.2.1 Methods

3.2.2.1.1 *Participants*

Overall 26 subjects participated in this EEG study. The data of one subject was excluded from all subsequent analyses leaving 25 participants (age range 19-31; mean age 25.44 (SEM .65); 13 females) for the final analysis.

3.2.2.1.2 *Stimuli*

The experimental stimuli used as standard and deviant were the same four VCV-syllables as in Experiment 3 such as [edi], [egi], [eni] and [ezi] (see Figure 13) which are all nonwords in German. The construction and controlled parameters have been described in section 3.2.1.1 Stimuli of Experiment 3, as well as in 2.2 Stimuli in General Methods).

Importantly, acoustic differences between the stimuli were minimized. Pitch and intensity measures were controlled for as best possible. The differences between the stimuli

are seen in Figure 17 illustrating the acoustic differences between the contrasts of each set of VCV-syllables.

Here, it is important to remember that the consonantal onset varied slightly between the [PLOSIVE] consonants ([d], [g]) and the [NASAL] [n] and [STRIDENT] [z]. The place and manner-of-articulation information of the [n] and [z] begin right after the offset of the [e]. However, the two plosives, [g] and [d] start with a period of silence after the offset of [e]. The closure duration lasts for 20 ms before the burst onset; consequently, the onset of the place and manner information of the two [PLOSIVE] sounds begins slightly later compared to the [n] and [z]. We were aware that this may add to latency differences in the MMN response for the conditions that contrast [NASAL] and [PLOSIVE] consonants. This was, however, considered in the parametrization of the latency measures of the MMN effects.

Additionally, all four sounds were evaluated for their word-medial frequencies as well as their phonotactic probabilities (see Table 8).

3.2.2.1.3 *Experimental Procedure*

In each experimental session, two of the four nonwords were combined pair-wise, with each nonword serving as a standard as well as a deviant, resulting in six blocks (Table 9) of a classical oddball paradigm. During the study, 680 standards (85%) and 120 deviants (15%) were presented for each of the six sequences.

3.2.2.1.4 *Data Acquisition and Analysis*

For each participant, the averaged MMN responses contained at least 77 % accepted deviant trials in each condition. The difference waveform (iMMN) was obtained by subtracting the ERP response of the standard from that of the equivalent deviant stimulus, e.g. [eni] as deviant minus [eni] as standard. Based on visual inspection of the grand-average waveform, the amplitude and latency of the MMN waveform of every individual condition was used and was expected around 150-250 ms after the 1st vowel offset (change onset), that is, in a time window around 320-450 ms.

For statistical analysis, the MMN peak latency and the mean amplitude of the MMN were used as dependant variables. The analyses were restricted to paired comparisons and *t*-tests because more complex factorial designs are not fully balanced. Full balancing was impossible for the present study where different featural dimensions varied and other possible confounds (physical differences between stimuli, number of conditions to signal to noise relationship, etc.) had to be avoided. The comparisons were restricted to three pairs of inversion with an equal acoustic change. Paired *t*-tests (*t*-test, α -level = .05) for the pairs of

inversion MANNER CHANGE ASYMMETRY ($[n]_{/d/} \sim [d]_{/n/}$), and MANNER CHANGE SYMMETRY ($[z]_{/n/} \sim [n]_{/z/}$) as control condition, as well as the pair of inversion MANNER & PLACE CHANGE SYMMETRY ($[g]_{/n/} \sim [n]_{/g/}$) were calculated.

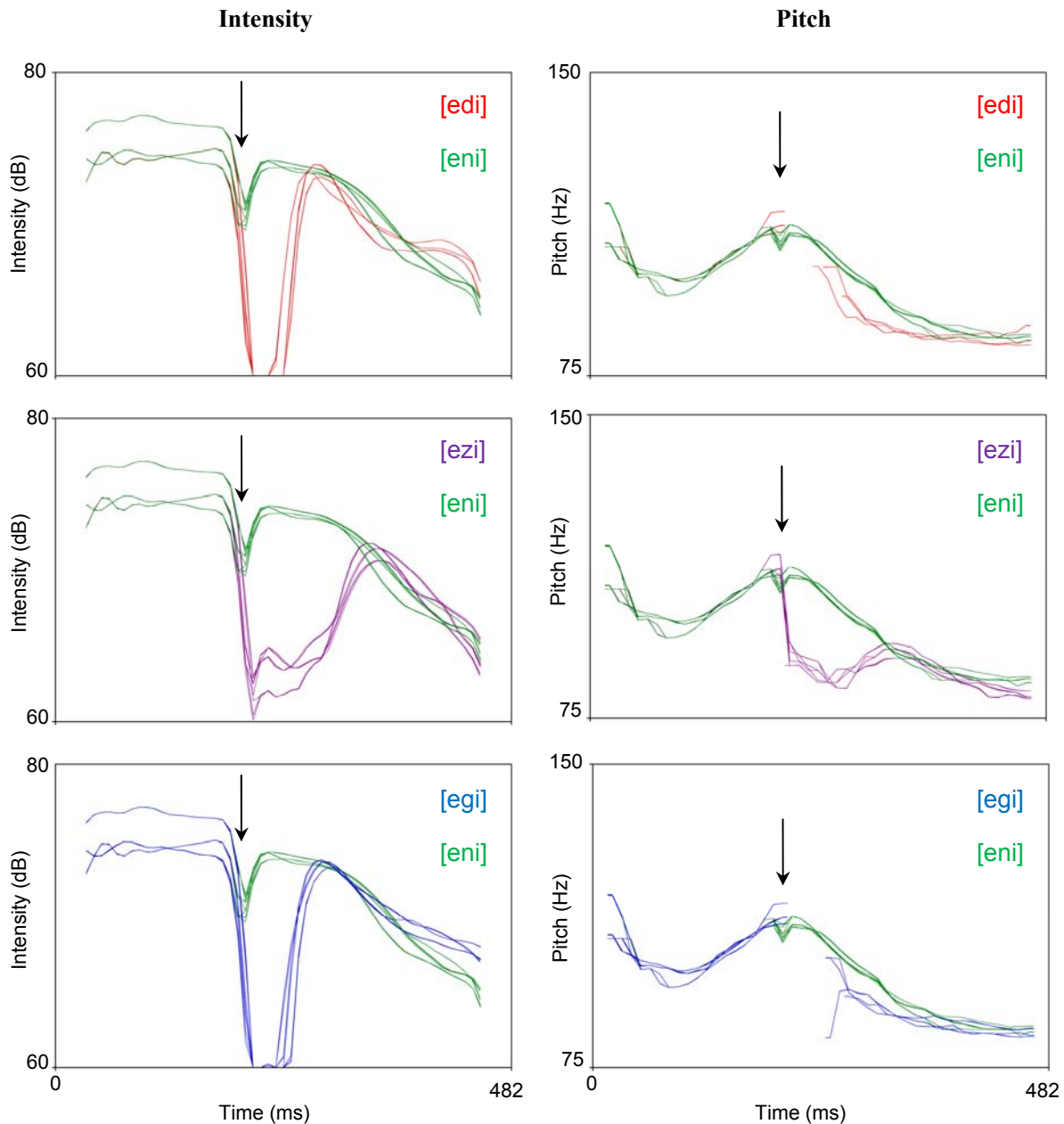


Figure 17: Pitch and intensity contours of the VCV-syllables used in this study: [eni] (green), [ezi] (violet), [edi] (red) and [egi] (blue). Three variants for each VCV-syllable were used, each set of a syllable showed comparable variation in pitch and intensity within each sound category. All three contrasts the [NASAL] ~ [PLOSIVE] contrasts [eni] ~ [edi] and [eni] ~ [egi] and the [NASAL] ~ [STRIDENT] contrast [eni] ~ [ezi] show substantial differences in intensity and pitch. Note, that the three [e:] sounds are the same across all four VCV-syllables. The arrows mark the onsets of the critical consonants.

3.2.2.2 Results

A clear MMN in the expected time window of the grand-average difference waveforms was observed in all six experimental conditions. The grand-average difference waveforms of all three contrasts can be seen in Figure 18. Figure 19 shows the mean MMN peak amplitudes for all three contrasts as bar graphs (see also Appendices for detailed figures of the grand-average waveforms of standards, deviants and difference waveforms including their MMN maps in Figures 5.1.6a-c).

The MMN amplitudes of the bi-directionally conflicting manner features MANNER CHANGE SYMMETRY ($[n]_{/z/} \sim [z]_{/n/}$) did not differ from each other ($t(24) = -0.54; p > .1; n.s.$). This was expected, when having manner conflicts in both directions of presentation.

The MMNs for the asymmetric manner contrast MANNER CHANGE ASYMMETRY revealed significant effects ($[d]_{/n/} \sim [n]_{/d/}$: $t(24) = 3.05; p < .006$). As predicted by the FUL-model, the asymmetric manner pair showed a larger MMN in the conflict condition, when the manner feature [PLOSIVE] of the deviant [d] mapped onto the pre-activated specified feature [NASAL] of the standard /n/. A reduced MMN amplitude was found in the reversed condition $[n]_{/d/}$, when the [NASAL] deviant mapped onto the underspecified [PLOSIVE] standard. The last contrast MANNER & PLACE CHANGE SYMMETRY ($[g]_{/n/} \sim [n]_{/g/}$), also showed significant differences ($[g]_{/n/} \sim [n]_{/g/}$: $t(24) = 3.30; p < .003$). This contrast had a manner conflict with [NASAL] in one direction ($[g]_{/n/}$) and a place conflict with [DORSAL] in the other ($[n]_{/g/}$). The manner conflict ($[g]_{/n/}$) revealed a significantly larger MMN amplitude than the place conflict.

Additionally, the differences between the contrasts were calculated and compared to each other to see whether there is a systematic difference between the contrasts. A significant difference between the MANNER CHANGE SYMMETRY ([NASAL] \sim [STRIDENT]) and the MANNER CHANGE ASYMMETRY ([NASAL] \sim [PLOSIVE]) contrast was found ($t(24) = 2.28; p < .032$), as well as a trend of a difference between the MANNER & PLACE CHANGE SYMMETRY contrast ([NASAL, CORONAL] \sim [PLOSIVE, DORSAL]) and MANNER CHANGE SYMMETRY contrast ([NASAL] \sim [STRIDENT]): ($t(24) = 1.83; p < .080$). No difference was seen between the MANNER CHANGE ASYMMETRY contrast and the MANNER & PLACE CHANGE SYMMETRY contrast ($t(24) = -0.41; p > .1; n.s.$; also see Appendices 5.2.2 Statistical Table 17a for detailed statistical analysis).

For consistency reasons across all experiments the same analysis was calculated again for condition-related mean MMN amplitudes for Fz and Cz. The mean amplitude was measured separately for each grand-averaged condition, resulting in six time-windows. This analysis showed similar results as for the individual mean MMN amplitudes (see Appendices 5.2.2 Statistical Table 17a for statistical analysis).

Since the stimuli had natural timing differences in the onsets between the [PLOSIVE] sounds and between the [NASAL] and [STRIDENT] sounds (see methods, 3.2.2.2 Stimuli), which might have influenced the MMN latency responses for the two [NASAL] ~ [PLOSIVE] contrasts $[d]_{/n/} \sim [n]_{/d/}$, and $[g]_{/n/} \sim [n]_{/g/}$, further latency analyses were conducted. Indeed, significant latency differences were found between the asymmetric manner contrast MANNER CHANGE ASYMMETRY ($[d]_{/n/} \sim [n]_{/d/}$: $t(24) = -2.62$; $p < .015$) and the MANNER & PLACE CHANGE SYMMETRY ($[g]_{/n/} \sim [n]_{/g/}$): $t(24) = -3.12$; $p < .005$). The [NASAL] and [STRIDENT] contrast did not differ in latency ($[n]_{/z/} \sim [z]_{/n/}$: $t(24) = 0.13$; $p > .1$; n.s.) which corresponds to the fact that they did not differ in their onset timing (for detailed statistical analysis see Appendices 5.2.2 Statistical Table 17b).

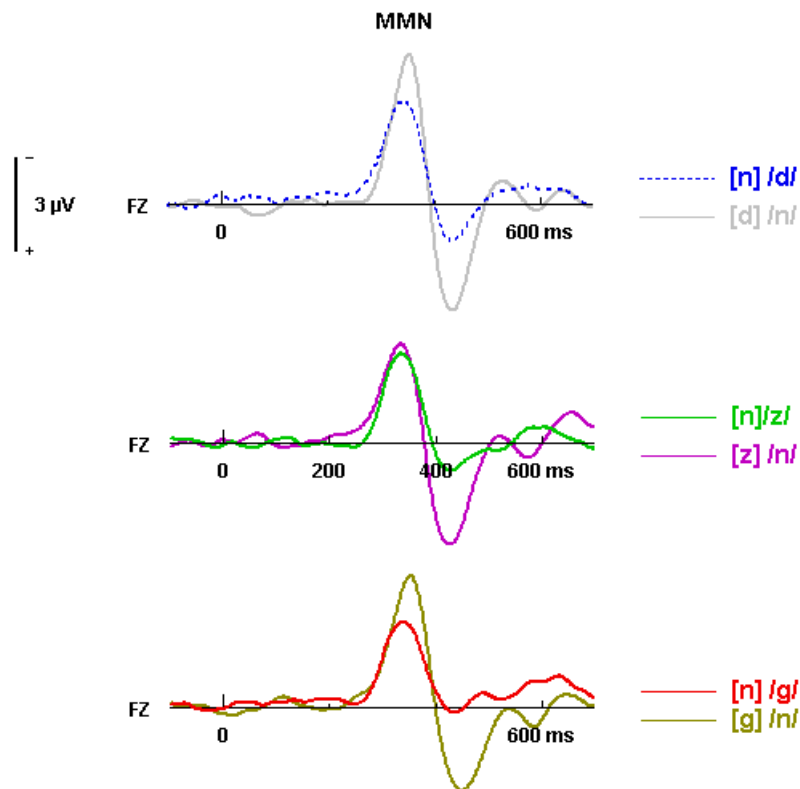


Figure 18: Grand-average waveforms of the frontal electrode (Fz) for all six conditions, the MANNER CHANGE ASYMMETRY Pair of Inversion ($[d]_{/n/}$ (grey) & $[n]_{/d/}$ (blue, dashed)), the MANNER CHANGE SYMMETRY pair ($[z]_{/n/}$ (violet) & $[n]_{/z/}$ (green)), and the MANNER & PLACE CHANGE SYMMETRY pair ($[g]_{/n/}$ (olive) & $[n]_{/g/}$ (red)). The only non-conflicting condition is dashed, all other conflicting conditions are solid lines. a), b), and c) refer to the three contrasts. All presented MMNs are re-referenced against their right mastoids.

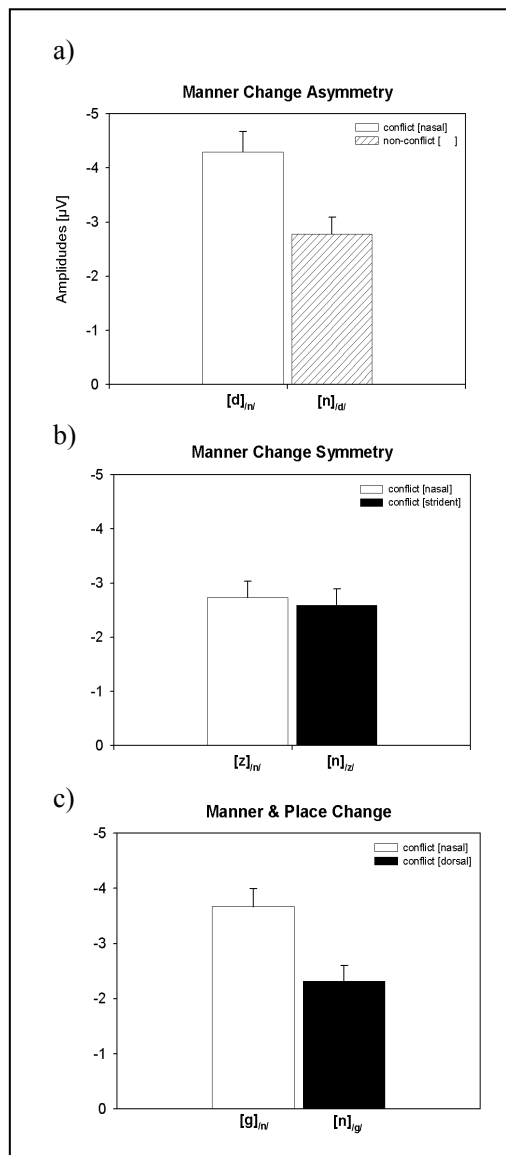


Figure 19: MMN mean peak amplitudes of all six conditions, the *Pair of Inversion MANNER CHANGE ASYMMETRY*, the control condition *MANNER CHANGE SYMMETRY*, and the pair *MANNER & PLACE CHANGE SYMMETRY*. a), b), and c) refer to the three contrasts. Error markings on top of the bars represent the standard error of the mean (SEM).

P300 Effect

Again there were large positivities seen in the ERP waveforms (Figure 18; Appendices Figure 5.1.6a-c), which were largest at Fz and Cz, and rather small at Pz, which could reflect a P3a having more frontal distributions (Squires et al., 1975; Knight et al., 1989; Katayama & Polich, 1998). As in Experiment 3, these P3a-like components were parametrized in a similar way as was done for the MMN and another statistical analysis was carried out. The mean amplitude of the positivity at the Fz electrode was used as a dependant variable and was measured over a 50 ms time window from positivity onset to offset (400-450 ms). The statistical analyses were restricted to paired comparisons and *t*-tests with the three pairs of inversion. Paired *t*-tests (*t*-Test, α -level = .05) for the pairs of inversion *MANNER CHANGE ASYMMETRY* ([n]_{dV} ~ [d]_{nV}), and *MANNER CHANGE SYMMETRY* ([z]_{nV} ~ [n]_{zV}), as well as the

pair of inversion MANNER & PLACE CHANGE SYMMETRY ($[g]_{/n/} \sim [n]_{/g/}$) were calculated. Amplitude values of the difference waves were again tested against zero in two-tailed t -tests separately for each condition.

This time, five of the six conditions revealed significant positivities ($[n]_{/z/}$: $t(24) = 2.25$; $p < .034$; $[z]_{/n/}$: $t(24) = 7.09$; $p < .000$; $[n]_{/d/}$: $t(24) = 2.46$; $p < .022$; $[d]_{/g/}$: $t(24) = 7.14$; $p < .000$; $[g]_{/n/}$: $t(24) = 4.98$; $p < .000$), only the condition $[n]_{/g/}$ was not significantly positive ($t(24) = 0.21$; $p > .1$; n.s.) in the P3a time window. Additionally all three contrasts differed significantly in their reversals: the MANNER CHANGE ASYMMETRY ($[n]_{/d/} \sim [d]_{/n/}$: $t(24) = -4.25$; $p < .000$), the pair MANNER CHANGE SYMMETRY ($[z]_{/n/} \sim [n]_{/z/}$: $t(24) = -4.53$; $p < .000$) as well as the pair MANNER & PLACE CHANGE SYMMETRY ($[g]_{/n/} \sim [n]_{/g/}$: $t(24) = -3.84$; $p < .001$). Whenever the [NASAL] was the standard or rather when a non-sonorant was the deviant in the case $[d]_{/n/}$, $[z]_{/n/}$ and $[g]_{/n/}$ the P3a amplitude was larger than the reversed case when the deviant was a sonorant (for detailed statistical analysis see Appendices 5.2.2 Statistical Table 17c).

3.2.2.3 Discussion of Experiment 4

The current study examined the fine structure of speech sounds during speech perception, by means of event-related brain potentials during the classic MMN paradigm. We compared the brain responses to sounds varying in different featural properties. In two contrasts, the comparison focused on change of manner features [NASAL], [PLOSIVE], and [STRIDENT], keeping place-of-articulation constant. The specific contrasts were (a) [NASAL] \sim [PLOSIVE] and (b) [NASAL] \sim [STRIDENT]. In the third contrast (c), we compared sounds differing in both manner and place of articulation: [NASAL] and [CORONAL] \sim [PLOSIVE] and [DORSAL]. In an oddball paradigm, one expects the brain responses to reflect the degree of difference between the standard and deviant sounds. Thus, the similarity or dissimilarity between the two would provide us with a means for measuring the degree of contrast between various sounds.

According to FUL's hypothesis, in contrast (a) an asymmetric MMN pattern was expected in the reversal of the [PLOSIVE] and [NASAL] manner of articulation, each serving as standard and deviant. In contrast (b), we assumed symmetric MMN patterns having conflicting conditions for both, [STRIDENT] and [NASAL] manner of articulation. The third contrast (c) is more challenging, since the sounds create conflicts in both directions of presentation, but for different types of features. Given that we have two different types of conflicts, *manner* and *place*, the question we raised was whether the MMN responses for each type of conflict were the same or not. If *place* and *manner* conflict evoke similar MMN

responses, the pattern would be the same as the contrast in (b) where the conflict in both directions involved manner features. Otherwise, we should see a difference.

In models where information of sound structure is fully stored (Bybee, 2001; Johnson, 2004; Norris, 1994; Norris & McQueen, 2008), the prediction would be that differences between standard and deviant stimuli would be reflected in MMN brain responses on the basis of acoustic discrepancies or statistical probabilities. Models where contextual experience plays a major role in dealing with phonological deviance (Gaskell & Marslen-Wilson, 1996; 1998), MMN differences would need to be explained by contextual properties. We discuss the results in light of our predictions as well as in terms of other hypotheses.

The results for contrasts (a) and (b) are fully in line with the predictions of the FUL model. We find asymmetric MMN activation patterns for our conflict vs. non-conflict manner contrast ($[n]_{/d/} \sim [d]_{/n/}$) and symmetric MMN amplitudes for the two-way manner conflict ($[n]_{/z/} \sim [z]_{/n/}$) and a significant difference between the two conditions, suggesting different processes between the asymmetric and the symmetric pair. The predictions laid out in Table 9 for the first two contrasts appear to hold. Before we turn to the third contrast, let us consider alternative explanations.

Can one account for these results purely based on the acoustic characteristics of these stimuli? The consonantal sounds used in the experiment — [d], [g], [n], [z] — differ quite substantially in their broad acoustic properties, particularly pitch and intensity (Figure 17). For example [n] is a sonorant while [d] is an obstruent, and consequently the differences in MMN pattern in contrast (a) could indeed be purely an acoustic difference. One could argue that a strong acoustic cue would cause a higher MMN. However, in our results the MMN amplitude of the deviant [d] compared to the standard /n/ is higher than the deviant [n] compared to standard /d/. For an acoustic explanation to hold for our MMN amplitude asymmetries in this contrast, one needs to claim that the release of [d] is acoustically stronger than that of [n], which is not evident from our stimuli. There is more reason to doubt a purely acoustic explanation if we look at the contrast (b) where the nasal [n] and the strident [z] are acoustically (pitch and intensity) rather different but we find no differences in their MMN activation pattern. Again, if it were due to the acoustics alone, it would be difficult to argue that the strident and nasal stimuli are equal in their acoustic strength. We believe that the equal MMNs in contrast (b) along with the asymmetric MMN in contrast (a) supports our claim that [NASAL] and [STRIDENT] are both specified in the mental representation and that these features conflict each other in the same way, whereas [PLOSIVE] is different.

Contrary to the MMN amplitude differences, the latency differences, however, between the [NASAL] and [PLOSIVE] sounds in contrast (a) (with [n]_{/d/} latency being slightly earlier) do seem to be caused by differences in duration of the consonantal information's onset. Recall that the plosives had an extra 20 ms silent closure duration before the burst onset. When [n] is the deviant with minimally earlier onset of the crucial information compared to /d/ as standard, then it affects the MMN in exactly this way, namely peaking earlier. Thus, these latency differences seem to be of acoustic nature triggered by the way the stimuli were produced.

Our asymmetric results cannot be solely explained through other well-known factors influencing the MMN, such as pitch and other acoustic differences (for review see Näätänen, 2001). We controlled for these factors, also by using the identity MMN approach, where the analysis is based on a comparison of identical stimuli, in their roles as standard and deviant (Eulitz & Lahiri, 2004; Pulvermüller & Styrov, 2006a).

Among other possible parameters, which could affect our results, a possible influence of phonotactic probabilities might be of relevance. Bonte et al. (2005) reported higher MMN amplitudes for nonwords with high phonotactic probability (*notsel*) as compared to the low probability condition (*notkel* and *notfel*). The distributional probabilities of [VC] sequences in our experiment were lowest for the cluster with [z], while all other stimuli were fairly similar (see 3.2.1.1 Stimuli & Table 8). According to the results of Bonte et al. (2005), the prediction for the present study would be a higher MMN amplitude for [eni], [edi] and [egi] compared to [ezi]. However, the pattern of MMN differences found in the present study differed considerably. The conditions with the similar phonotactic probabilities show large MMN differences ([edi] < [eni]; [egi] < [eni]), whereas the conditions with a larger difference in phonotactic probabilities showed no MMN difference ([ezi] = [eni]). Thus, contrary to Bonte et al. (2005), no MMN amplitude difference between our high vs. low phonotactic probability condition was found.

Furthermore, our findings cannot be explained by individual sound frequency effects of our stimuli. The intervocalic frequency turned out to be highest for [n] and [g], slightly lower for [d] and lowest for [z] (see Table 8). Again, one could argue that a higher sound frequency deviant would elicit a higher MMN response; however, in our results the MMN amplitude of the deviant [d] and [g] as compared to the standard /n/ is increased compared to the reversed conditions. Additionally, the largest sound frequency difference is again between [n] and [z], but here we find equal MMN amplitudes. These results show a pattern, which cannot be explained by sound frequency effects or phonotactic probability influences.

With respect to more probabilistic connectionist models such as contextual integration (Gaskell et al., 1995, Gaskell & Snoeren, 2008; Gaskell & Marslen-Wilson, 1996, 1997), these crucially depend on notions of assimilation and recognition of variant stimuli, which have changed due to neighboring contexts. We examined intervocalic medial consonants where the vocalic contexts were deliberately kept constant. The stimuli were also cross-spliced to maintain as much acoustic similarity as possible. Thus, context or experience could not have played a role in explaining these results.

We now turn to our third contrast where we examined consonants that conflict in manner of articulation in one direction and place of articulation in the opposite direction. By varying both manner and place features in the same contrast, we wanted to explore to what extent MMN responses to two conflicting features in different feature domains would interact. Our results showed an asymmetric MMN pattern, an increased amplitude for the manner conflict compared to the place conflict. That is, the plosive deviant with nasal standard triggered a higher MMN amplitude than the coronal deviant to the dorsal standard. This contrast was different from the symmetric [NASAL] ~ [STRIDENT] contrast by trend, also suggesting that there might be something further going on. One could speculate that manner of articulation might play a stronger role in the extraction of feature information. First, the perceptual system extracts manner of articulation, deciding whether a sound is constricted, completely stopped, or nasal, and so on, before it extracts information about where a sound is articulated.

If we argue on the basis of acoustic characteristics, one would need to claim that the sound with more energy and higher pitch triggers higher MMNs, which could indeed be the case for our condition (c), the [NASAL] ~ [PLOSIVE] contrast. However, the acoustic differences between [n] and [g] do not support this asymmetric MMN pattern, rather showing the reverse.

One cannot argue that place differences do not trigger MMN differences because other studies have shown differences in MMN responses between plosive sounds differing only in place of articulation (see Experiment 3; Dehaene-Lambertz, 1997; Maiste et al., 1995; Shafer, et al., 2004; Walter & Haquard, 2004), despite their very close physical pitch and intensity attributes. All these studies show that the coronal [d] deviant triggers a higher MMN compared to the [g] or [b] non-coronal deviants. The explanations for these differences vary. For instance, Dehaene-Lambertz (1997) and Maiste et al. (1995) argue in favor of formant transition differences, while Cornell et al. (2009; Experiment 3) and Walter and Haquard (2004) favor an underspecification representational hypothesis. Be that as it may, what is crucial here is that the asymmetric pattern for [g] ~ [n] could reflect a difference in class structure between the features. From contrast (a) we know that [NASAL] ~

[PLOSIVE] asymmetry exists. From other research we know that coronal ~ non-coronal asymmetry exists within plosives and nasals (Walter & Haquard, 2004; Experiment 3). Consequently, the pattern we see for contrast (c) could be hierarchical, where the [NASAL] ~ [PLOSIVE] contrast triggers a higher MMN than a [DORSAL] ~ [CORONAL] place contrast. Nevertheless, to ensure that the MMN is sensitive to hierarchical effects in multiple feature conflicts, further research is necessary, comparing more place and manner features across dimensions with multiple conflicts and different asymmetries.

P300 – Automatic Attention Switch

In Experiment 4, the results revealed the presence of robust positive components immediately following the MMN in five of the six conditions, with larger amplitudes for the non-sonorant sounds [z], [d], and [g] as deviants compared to the sonorant [n] as deviant. This effect was again more evident at the Fz and Cz and as already described in Experiment 3 this positivity could very well be related to the P3a as it has a frontocentral scalp distribution (Squires et al., 1975; Knight et al., 1989; Katayama & Polich, 1998). It is thought to reflect a switch of attention toward the irrelevant deviant when the deviant is considerably different from the context it is presented in (Sams et al., 1985a; Escera et al., 2000). A significant change in the background ‘noise’ will cause a shift in attention i.e., an orienting response, which is reflected in the P3a.

This oddball manipulation of deviant / standard sound stimulus context seemed to have produced a task environment in which the deviant sounds [z], [d], and [g] quickly engaged focal attention in a manner similar to the one observed for highly ‘novel’ stimuli. The frontal maximum deviant P3 of the present study can, therefore, be considered a manifestation of the P3a, which may be the same as the novelty P300 (Squires et al., 1975). What is considered novel and / or salient depends also on the context in which the eliciting event is encountered, even to familiar stimuli (Friedman et al., 2001).

It could be argued that if a context is defined primarily by a non-sonorant deviant and a sonorant standard discrimination, attentional redirection to the non-sonorant would occur by means of frontal lobe activation that may underlie P3a components. In the reversed case, a sonorant deviant and non-sonorant standard discrimination however, would diminish this effect because the stimulus change would not demand as much attentional effort so that the deviant sonorant stimulus could more easily be ignored. The strident [z], as well as the plosives [d] and [g] as deviants seem to mismatch stronger, registering a more robust change compared to the reversed conditions when the deviant is a more harmonious sound, as is the case of [n]. The non-sonorant sounds seem to attract more attentional resources being more salient. It is suggested, that when the acoustic qualities of sounds differ radically from the qualities of the background stimuli (standards) – in our study perhaps the change from

[NASAL] to [STRIDENT] or [PLOSIVE] – the P3s effect reflects the orienting response in the first place with no distinct differentiation of the sound nature per se.

As already mentioned, the MMN often is followed by the P3a, even to small changes in the to-be-ignored stimulus sequences. This response might indicate the actual orienting of attention to an MMN-eliciting sound change. If the event is sufficiently deviant, as it seems the case for our non-sonorant sounds, the MMN is followed by the P3a. A process has been triggered that is active after the deviant has been detected by the brain, as opposed to the detection of deviant events per se, which seems to be realized at a pre-attentive level reflected by the MMN.

In conclusion, this present study provides electrophysiological evidence for the notion of specified and underspecified cortical representations for manner of articulation, finding asymmetric mappings of features extracted from the acoustic speech signal onto their underlying representations for different manner features. Our results cover new ground by contrasting manner of articulation in medial consonantal sounds, generalizing the previous findings of asymmetrical activation patterns to other feature dimensions, such as [NASAL], [STRIDENT], and [PLOSIVE]. If place of articulation is held constant, our MMN asymmetries between the [NASAL] and [PLOSIVE] sounds in the conflict - non-conflict condition strongly speak for the underspecification of the feature [PLOSIVE]. However, if consonants conflict in manner in one direction and place in another, we find enhanced MMN for manner. The results of the last condition seem to support a hierarchy in using featural dimensions during speech perception, with manner of articulation being superior over place in the extraction of feature information. Among all possible reasons for our MMN differences discussed above, the fine structure of phonological underspecified abstract representations in the mental lexicon explains the pattern of results best. Breaking new grounds by looking at other feature dimensions such as manner of articulation this study provides further support for the suggestion that the brain refers to more abstract underspecified phonological representations based on phonological principles which are not contextually dependant and do not reflect a one-to-one relation between the speech signal and their mental representation. If the sounds were intrinsic wholes, we should not expect any differences or only differences based on frequency or contextual effects.

3.2.3 Experiment 5: Mental Representations of Place- and Manner-of-Articulation Features in Different Speech Sounds

The goal of this current study is to further examine the phonological fine structure of mental representations of manner features, as well as place features and compare these two feature dimensions with different consonantal sounds.

In the earlier studies (Experiments 1-4), fine grained phonological contrasts were examined in vowels and in stop consonants for the place features [CORONAL] and [DORSAL] (Eulitz & Lahiri, 2004; Experiments 1 & 3), as well as the manner features [PLOSIVE], [NASAL] and [STRIDENT] (Experiments 3 & 4), finding asymmetric activations between specified and underspecified feature representations. In particular, Experiment 3 found asymmetric activations between stop consonants for the place contrasts [CORONAL] and [DORSAL]. Experiment 4 could show MMN asymmetries in the [PLOSIVE] and [NASAL] contrast for coronal sounds.

One aim of this Experiment 5 was to extend these findings, further investigating the place, as well as the manner of articulation dimension with different speech sounds. We wanted to investigate whether the small MMN asymmetries in the place contrast of acoustically similar plosive sounds (Experiment 3) can be further generalized to acoustically comparable sounds such as nasals. At the same time we wanted to expand the place contrasts found for [CORONAL] and [DORSAL] stops, to [CORONAL] ~ [LABIAL] contrasts, both for nasals and plosives. Furthermore, it was of interest to replicate the asymmetries of the manner features [PLOSIVE] and [NASAL] for coronals of Experiment 4 and further generalize this contrast to labials.

We used the four consonantal sounds [n], [m], [d] and [b] embedded in meaningless VCV-syllables (e.g. *eni, emi, edi, ebi*) in eight experimental conditions. The sounds [n] and [m] are distinguished by the feature [NASAL] differing in their place of articulation, the former being [CORONAL] and the latter [LABIAL]. The sounds [d] and [b] are voiced stops distinguished by the feature [PLOSIVE] and also differ in their place of articulation. The plosive [d] shares [CORONAL] place with [n], whereas [b] shares its [LABIAL] place with [m] (In Table 10, the third column lists the surface features of the consonants in square brackets listed in the first column). It is suggested that not all of these sounds have the same status, in that there are asymmetric relationships between them.

Since one goal of this current study was to generalize the findings of the previous experiments (3 & 4) to other speech sounds, we contrasted the nasals [n] ~ [m] and the plosives [d] ~ [b] varying [CORONAL] ~ [LABIAL] place of articulation. As for our manner contrasts, we varied the coronal sounds [n] ~ [d] and the labial sounds [m] ~ [b] altering [NASAL] ~ [PLOSIVE]. Based on our hypothesis of underspecification and previous studies

(e.g. Eulitz & Lahiri, 2004; Walter & Haquard, 2004; see also Experiments 3 & 4), these contrasts should lead to asymmetric activations in all cases according to their assumed differences in mental representations (details of the representations of the various consonants are given in column 3 of Table 10, column 2 lists the features that are extracted from the signal; final column in Table 10 explains which features can conflict according to FUL). For the two place contrasts this means when [LABIAL] is extracted from the deviant signal, it does not conflict with the underlying representation since it maps onto an empty slot in the lexicon. However, this is not the case in the opposite direction - when [CORONAL] is extracted from the signal it conflicts with [LABIAL] since it is specified in the underlying representation. Therefore, we predict an asymmetry of MMNs within the reversal of consonantal contrasts changing place of articulation for both nasals and plosives, presented as standard and deviant. Larger MMN amplitudes are expected in conditions with a feature-conflict than in conditions with non-conflict conditions. More specifically, contrasting place of articulation of the nasal and plosive sounds, the [n]_{/m/} condition (labeling see Table 10) extracts the feature [CORONAL] of the deviant [n] and conflicts with [LABIAL] in the mental representation formed by the standard /m/. In the reversed condition, [m]_{/n/} when [LABIAL] from the deviant [m] is extracted and maps onto the pre-activated representation of the underspecified coronal standard /n/, a non-conflicting situation occurs. Consequently, the MMN amplitude difference between these acoustically equalized conditions [n]_{/m/} (conflict) ~ [m]_{/n/} (non-conflict) should be larger for the conflicting condition (or earlier in latency, following Eulitz & Lahiri, 2004), resulting in an asymmetric MMN pattern between the nasals. The same asymmetric pattern would be expected for the plosive sounds: an MMN amplitude difference between the acoustically equalized conditions [d]_{/b/} (conflict) ~ [b]_{/d/} (non-conflict), with a larger amplitude (or earlier latency) for the conflicting condition.

The predictions for the Manner contrasts are comparable when [NASAL] is extracted from the deviant signal, it does not conflict with the underlying representation since it maps onto an empty slot in the lexicon. For the opposite direction - when [PLOSIVE] is extracted from the signal it conflicts with [NASAL] which is specified in the underlying representation. Therefore we also predict asymmetries of MMNs within the reversal of consonantal contrasts for both coronal and labial sounds, presented as standard and deviant. An enhanced MMN amplitude (or earlier latency) is expected for our conflicting coronal contrast [d]_{/n/} as opposed to our non-conflicting coronal contrast [n]_{/d/}. Similarly, we assume to find an asymmetric MMN response pattern for our labial contrast: a larger amplitude (or earlier latency) for the conflicting condition [b]_{/m/}, compared to the non-conflicting condition [m]_{/b/} (for summary see Table 10).

Since the stimuli had natural timing differences in the onsets between the [PLOSIVE] and the [NASAL] sounds (see methods, stimuli section), which might have influenced the

MMN responses for the two [NASAL] ~ [PLOSIVE] contrasts [d]_{/n/} ~ [n]_{/d/}, and [m]_{/b/} ~ [b]_{/m/}, latencies were analyzed for all four contrasts. It was expected, that latency differences might occur due to these acoustic differences of the stimuli between the two [NASAL] ~ [PLOSIVE] manner contrasts. Furthermore, it was assumed that there would be no latency differences within the place contrasts due to acoustic differences, since their onset timing is the same between these contrasts. However, in line with previous work, one could further assume to find latency differences for the conflicting as opposed to the non-conflicting conditions. Eulitz and Lahiri (2004) found latency differences in their vowel pairs, an earlier latency for the conflict condition in their [CORONAL] ~ [DORSAL] contrast. Walter and Haquard (2004) could show a tendency of the same latency effect between their nasal and plosive sounds varying [CORONAL] and [LABIAL]. Thus, latency differences could either be due to acoustic differences (between [NASAL] ~ [PLOSIVE]) or due to the assumption of underspecification, having an earlier effect for the conflicting conditions.

It is important to realize that these predictions of asymmetric activation patterns laid out here are assumed due to the theory of underspecification in the mental representations (e.g. Lahiri & Reetz, 2002; 2010). Other models assuming that the core representation of a lexical form is fully specified (e.g. Bybee, 2001; Norris, 1994; Norris & McQueen, 2008; Gaskell & Marslen-Wilson, 1997; Gaskell et al., 1995), expect to find equal MMNs between all four contrasting consonantal sounds with varying place and manner features since it is assumed that all phonological information is fully stored. Further, models dealing with phonological variation in the speech stream are either based on experience and frequency of occurrence, or on contextually receptive information. Variation is compensated for by inference systems (e.g. Gaskell & Marslen-Wilson, 1996, 1998, 2001) or probabilistic computations (Norris & McQueen, 2008). They would presumably predict equal activation patterns, since our stimuli are intended in a way that context remains constant (they all have the same vowel contexts), or the activation patterns would need to mirror frequency and probabilistic effects.

Table 10: Predictions about the amount of feature conflict for the nonword stimuli used in the eight experimental conditions

Experimental conditions [Deviant]/Standard/	Features in the surface representation (extracted by the deviant)		Features in the lexical representation (activated by the standard)		Full specification hypotheses	Underspecification hypotheses	
	MoA	PoA	MoA	PoA			
Place change							
a)	[n] _{/m/}	[NAS]	[COR]	[NAS]	[LAB]	conflict	conflict with [LAB]
	[m] _{/n/}	[NAS]	[LAB]	[NAS]	[]	conflict	non-conflict []
b)	[d] _{/b/}	[PLO]	[COR]	[]	[LAB]	conflict	conflict with [LAB]
	[b] _{/d/}	[PLO]	[LAB]	[]	[]	conflict	non-conflict []
Manner change							
c)	[n] _{/d/}	[NAS]	[COR]	[]	[]	conflict	non-conflict []
	[d] _{/n/}	[PLO]	[COR]	[NAS]	[]	conflict	conflict with [NAS]
d)	[m] _{/b/}	[NAS]	[LAB]	[]	[LAB]	conflict	non-conflict []
	[b] _{/m/}	[PLO]	[LAB]	[NAS]	[LAB]	conflict	conflict with [NAS]

The standard-deviant contrasts are identical in each pair of inversion. Note, that for the place contrasts the manner of articulation is kept constant, for the manner contrasts the place of articulation is kept constant. The shaded areas highlight the non-conflicting conditions assuming underspecification. MoA = manner of articulation; PoA = place of articulation; [COR] = [CORONAL]; [LAB] = [LABIAL]; [NAS] = [NASAL]; [PLO] = [PLOSIVE]; [] = underspecified feature.

3.2.3.1 Methods

3.2.3.1.1 Participants

Twenty-seven students participated in this EEG study. The data of three subjects were excluded from all subsequent analyses due to excessive EOG artifacts or other noise. The remaining 24 participants (age range 20-26 years; mean age 23.13 (SEM .46); 50% female) went into the final analysis.

3.2.3.1.2 Stimuli

The experimental stimuli used as standard and deviant were four VCV-syllables such as [edi], [ebi], [eni] and [emi] (Figure 20). Note that the stimuli [edi] and [eni] are the same as in the other previous experiments (Experiment 3 & 4; see methods sections 3.2.1.2 Stimuli, 3.2.2.2 Stimuli, and 2.2 Stimuli in General Methods). The stimuli [emi] and [ebi] were

recorded separately from the other stimuli specifically for this experiment but with the same speaker. These stimuli were manipulated in exactly the same way as all other stimuli.

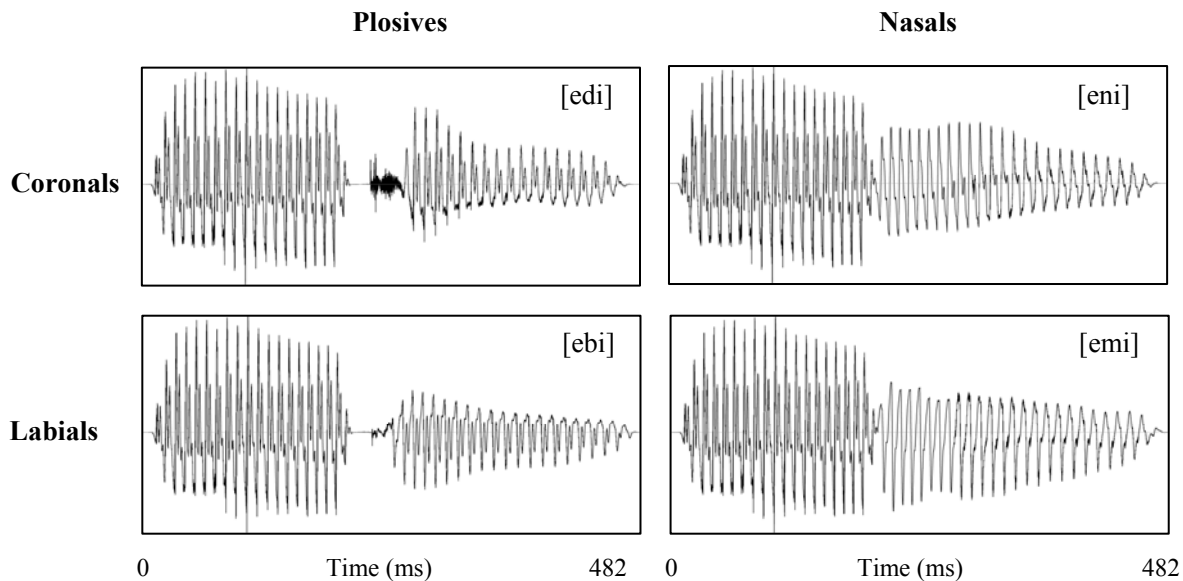


Figure 20: Oscillograms of the acoustic stimuli used in the experiment: [edi], [ebi], [eni] and [emi]. The standard and deviant stimuli were maximally matched for their acoustic properties in all conditions.

Again, three variants of each stimulus type with equivalent F0 frequencies were selected, as well as pitch and intensity measures were controlled for as best possible. Figure 21 illustrates that each set of VCV-syllables show either comparable variation in pitch and intensity for nasals and plosives, or larger variation in pitch and intensity between coronals and labials.

It is important to remember that the consonantal onsets vary between the [PLOSIVE] consonants ([d], [b]) and [NASAL] consonants ([n], [m]) due to the closure duration of the plosive sounds. We were aware that this might cause latency differences in the MMN response for the conditions that contrast [NASAL] and [PLOSIVE] consonants which has already been the case in Experiment 4 and measured them accordingly.

All stimuli were equalized for intensity (RMS). Since the RMS varied naturally between the experimental stimuli, the intensities were perceptually adjusted (see loudness rating in methods section of Experiment 3) with the following amplitudes: 74dB for the initial vowel [e], 68dB for [di] and 70dB for [ni]. The newly recorded sounds were adjusted following their sound category: 68 db for [bi] and [70 dB] for [mi] and subjectively rated by four phoneticians to be perceptually equal to the other sounds. A linear onset and offset ramp of 20 ms (linear fade in, linear fade out; Adobe Audition) was applied to the intensities in order to smooth the beginning and ends of our stimuli.

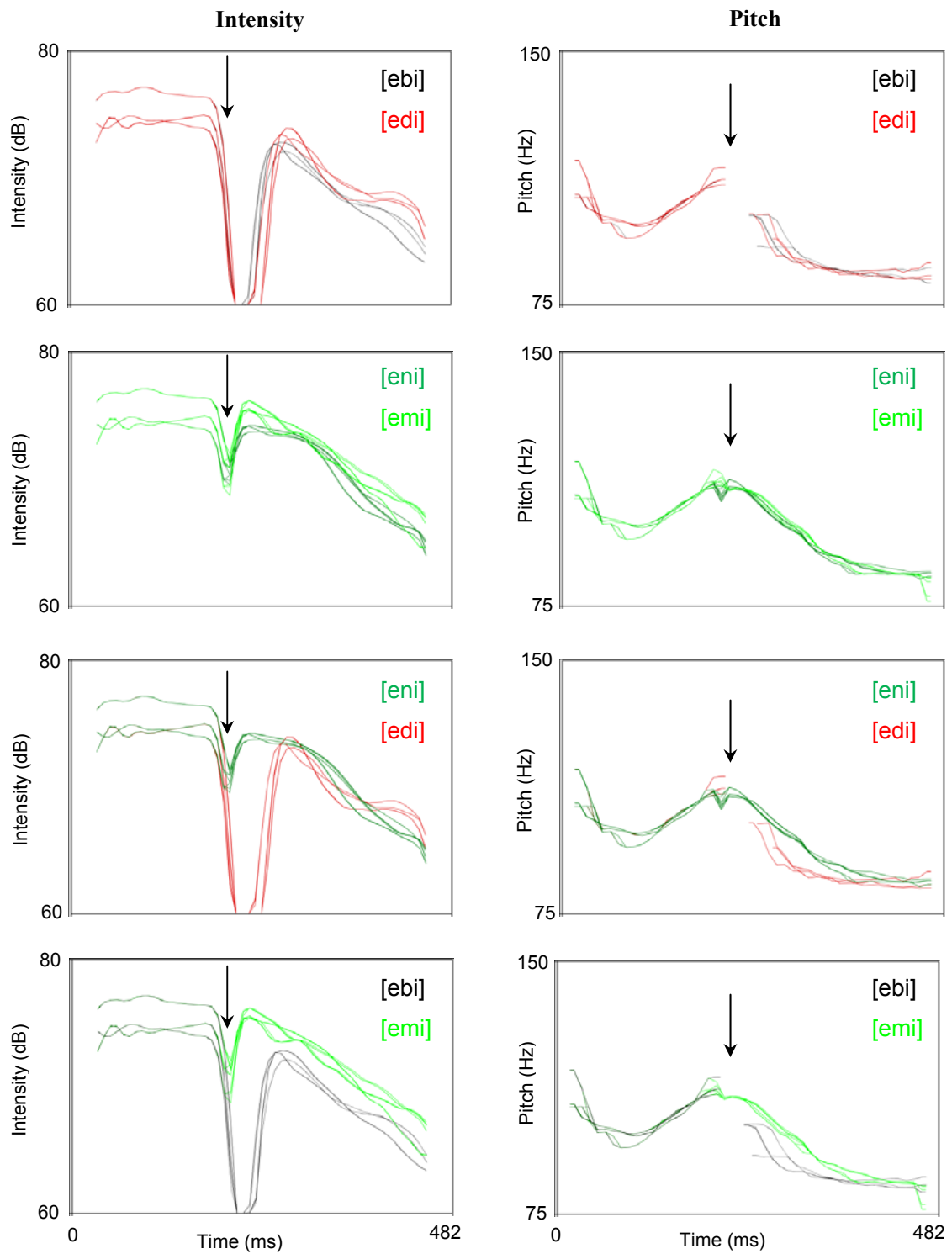


Figure 21: Pitch and intensity contours of the VCV-syllables used in this study: [eni] (green), [emi] (lime), [edi] (red) & [ebi] (black). Three variants for each VCV-syllable were used, each set of syllable showed comparable variation in pitch and intensity. The plosive place contrasts [eni] ~ [emi] and [edi] ~ [ebi] are very similar in pitch and intensity, the manner contrasts [eni] ~ [edi] & [ebi] ~ [emi] differ in pitch and intensity. Note, that the three [e:] sounds are the same across all four VCV-syllables. The arrows mark the onsets of the critical consonants.

All four sounds were again checked for their word-medial frequencies (Table 11). The consonantal sounds differ in their frequency counts in medial intervocalic position, with [m] and [b] being fairly similar ([m] log-value 1.83, [b] log-value 1.26), having lower frequencies than [d] (log-value 2.84), and [n] having the highest frequency (log-value 3.74) (second column in Table 11).

Again, distributional frequencies of the speech sounds (phonotactic probabilities), were checked for the first critical vowel - consonant ([VC]) syllables (third column in Table 11). The [VC] sequences were fairly similar to each other, [eb] with the highest probability, [en] being slightly lower, and [em] and [ed] having both equivalent lower probability values (log-values: [eb] 3.67 > [en] 2.81 > [ed] 2.25 = [em] 2.26). The frequencies of the critical consonants of words, which had consonants in the beginning and end, i.e. in the environment of C[VCV]C, differed minimally from each other (fourth column in Table 11).

Table 11: Log-values of the frequency counts of phonemes and phoneme sequences in the studied VCV-syllables:

nonword stimuli	V[C]V	[VC]V	C[VCV]C
[eni]	3.74	2.81	5.06
[emi]	1.83	2.26	4.70
[edi]	2.84	2.25	4.73
[ebi]	1.26	3.67	5.16

Frequency counts were weighed for word frequency and are based on the CELEX corpus. [...] indicates the phonemes for which counts are given in the table; C = consonant; V = vowel; e.g. C[VCV]C in the fourth column gives the frequency counts for /n/, /m/, /d/ and /b/ in an intervocalic position with further sounds before and after the vowels (e.g. words such as N[ome]n ‘noun’).

3.2.3.1.3 *Experimental Procedure*

In each experimental session the four nonwords were combined in two pairs, with each nonword serving as a standard as well as a deviant, resulting in eight blocks (Table 10) of a classical oddball paradigm. During the study, 680 standards (85%) and 120 deviants (15%) were presented for each of the eight blocks.

3.2.3.1.4 *Data Acquisition and Analysis*

For each participant, the averaged MMN responses contained at least 77 % accepted deviant trials in each condition. The MMN was obtained by subtracting the ERP response of the

standard from that of the equivalent deviant stimulus across blocks, e.g. [emi] as deviant minus [emi] as standard of the reversed block, resulting in the iMMN.

Based on visual inspection of the grand-average waveform, the maximum negative amplitude of the MMN difference waveform of every individual condition was used for quantification of the ERPs. The MMN was expected around 150-250 ms after the 1st vowel offset (change onset), that is, in a time window around 350-450 ms. For statistical analysis, the individual peak latency and mean amplitude of the MMN were used as dependant variables.

The statistical model was restricted to two repeated-measure ANOVAs for each mean amplitude and peak latency value for the place and manner contrasts with the factors FEATURE-TYPE (first ANOVA: place contrast: nasals vs. plosives; second ANOVA: manner contrast: coronals vs. labials) and LEVEL OF CONFLICT (first ANOVA: place contrast: conflict conditions [n]_{/m/} & [d]_{/b/} vs. non-conflict conditions [m]_{/n/} & [b]_{/d/}; second ANOVA: manner contrast: conflict conditions [d]_{/n/} & [b]_{/m/} vs. non-conflict conditions [n]_{/d/} & [m]_{/b/}). Only significant main effects or interactions are reported. In the case of significant interactions paired *t*-tests were used for post-hoc testing. The Greenhouse-Geisser adjustment was used where appropriate and the corrected *p*-values are reported together with the uncorrected degrees of freedom. Additionally, amplitude values of the difference waves were tested against zero in two-tailed *t*-tests separately for each condition to test for MMN existence.

3.2.3.2 Results

Significant MMNs in the expected time window of the grand average difference waveforms were observed in all eight experimental conditions. There was sizable variation in the amplitude. The grand-verage difference waveforms of all four contrasts can be seen in Figure 22. Figure 23 shows the individual MMN mean peak amplitudes and peak latencies of all four contrasts in bar graphs (see also Appendices for more detailed figures of the grand-average waveforms of standards, deviants and difference waveforms including their MMN maps in Figures 5.1.7a-d; for detailed statistical analysis see Appendices 5.2.2 Statistical Table 18a).

The ANOVA of the place contrasts (a) and (b) revealed only a main effect for FEATURE-TYPE in the amplitude measure ($F(1, 23) = 27.43; p < .000$), with larger amplitudes for the plosives. No further interaction or main effect of LEVEL OF CONFLICT was seen. Hypothesis-driven, paired comparisons of each FEATURE-TYPE was calculated, but no differences were found between the nasals ($F(1, 23) = 0.06; p > .1; n.s.$), nor the plosives ($F(1, 23) = 0.51; p > .1; n.s.$).

The ANOVA of the manner contrasts (c) and (d) however, revealed a significant interaction with FEATURE-TYPE x LEVEL OF CONFLICT of the MMN amplitude ($F(1, 23) = 6.09; p < .021$; Figure 21 & 22A), as well as a main effect of FEATURE-TYPE ($F(1, 23) = 12.36; p < .002$) and LEVEL OF CONFLICT ($F(1, 23) = 5.24; p < .032$). The coronal sounds elicited significantly larger amplitudes than the labial sounds. Paired comparisons of the FEATURE-TYPE showed a significant difference between the conflicting versus non-conflicting coronal sounds ($[d]_{/n/} \sim [n]_{/d/}$: $F(1, 23) = 12.64; p < .002$) (Figure 22 & 23 A). The conflicting condition $[d]_{/n/}$ triggered higher amplitudes as opposed to the reversed non-conflicting condition $[n]_{/d/}$. The labial contrast did not differ from each other ($[b]_{/m/} \sim [m]_{/b/}$: $F(1, 23) = 0.01; p > .1$; n.s.). Comparisons of the LEVEL OF CONFLICT demonstrated only a difference for the two conflicting conditions ($[d]_{/n/} \sim [b]_{/m/}$: $F(1, 23) = 11.87; p < .002$), but not for the two non-conflicting conditions ($[n]_{/d/} \sim [m]_{/b/}$: $F(1, 23) = 0.85; p > .1$; n.s.).

The same amplitude analysis was calculated for condition-related mean MMN amplitudes for Fz and Cz for consistency reasons across all experiments. This condition-related analysis was measured separately for each condition, resulting in four grand-average time windows of each condition. The analysis showed the same results as for the individual mean MMN amplitudes (see detailed statistical analysis in Appendices 5.2.2 Statistical Table 18a).

The same ANOVAs were calculated for the MMN peak latencies as was done for the amplitude measures. The ANOVA of the place contrasts revealed two main effects for FEATURE-TYPE ($F(1, 23) = 11.28; p < .003$) and for LEVEL OF CONFLICT ($F(1, 23) = 19.58; p < .000$) in the latency measure of the MMN (Figure 22 & 23 B). Hypothesis driven paired comparisons of the FEATURE-TYPE illustrated a significant difference between the nasals ($[m]_{/n/} \sim [n]_{/m/}$: $F(1, 23) = 5.92; p < .023$) and the plosives ($[b]_{/d/} \sim [d]_{/b/}$: $F(1, 23) = 12.08; p < .002$), the nasals being earlier. The comparison within the LEVEL OF CONFLICT further showed a difference between the conflicting conditions ($[n]_{/m/} \sim [d]_{/b/}$: $F(1, 23) = 10.85; p < .003$) and the non-conflicting conditions ($[m]_{/n/} \sim [b]_{/d/}$: $F(1, 23) = 8.93; p < .007$) with the conflicting conditions showing an earlier latency than the non-conflicting conditions. The ANOVA of the manner contrasts did not show any latency effects (all $F_s < 1$; for detailed statistical analysis see Appendices 5.2.2 Statistical Table 18b).

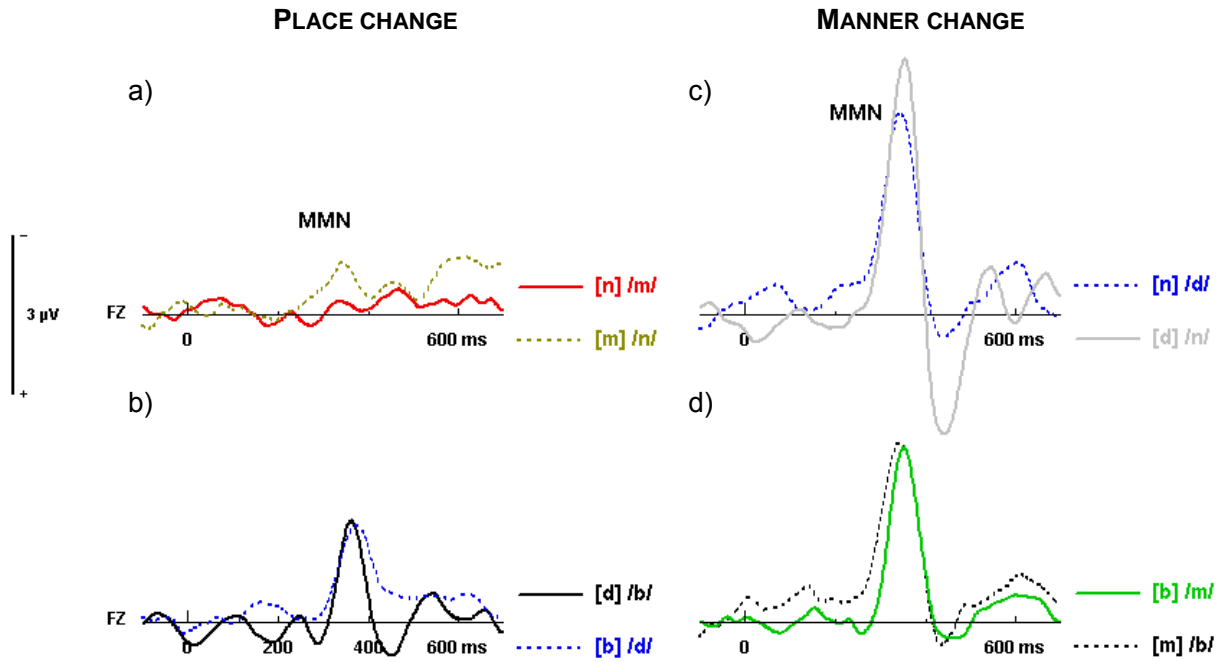


Figure 22: Grand-average waveforms of the frontal electrode (Fz) for all eight conditions, the PLACE contrasts with nasals ([n]_{/m/} ~ [m]_{/n/}) and plosives ([d]_{/b/} ~ [b]_{/d/}), and the MANNER contrasts with coronals ([d]_{/n/} ~ [n]_{/d/}), and labials ([b]_{/m/} ~ [m]_{/b/}). The non-conflicting conditions are dashed, all conflicting conditions are shown in solid lines. a) & b) refer to the two place contrasts, c) and d) refer to the two manner contrasts. All presented MMNs are re-referenced against their right mastoids.

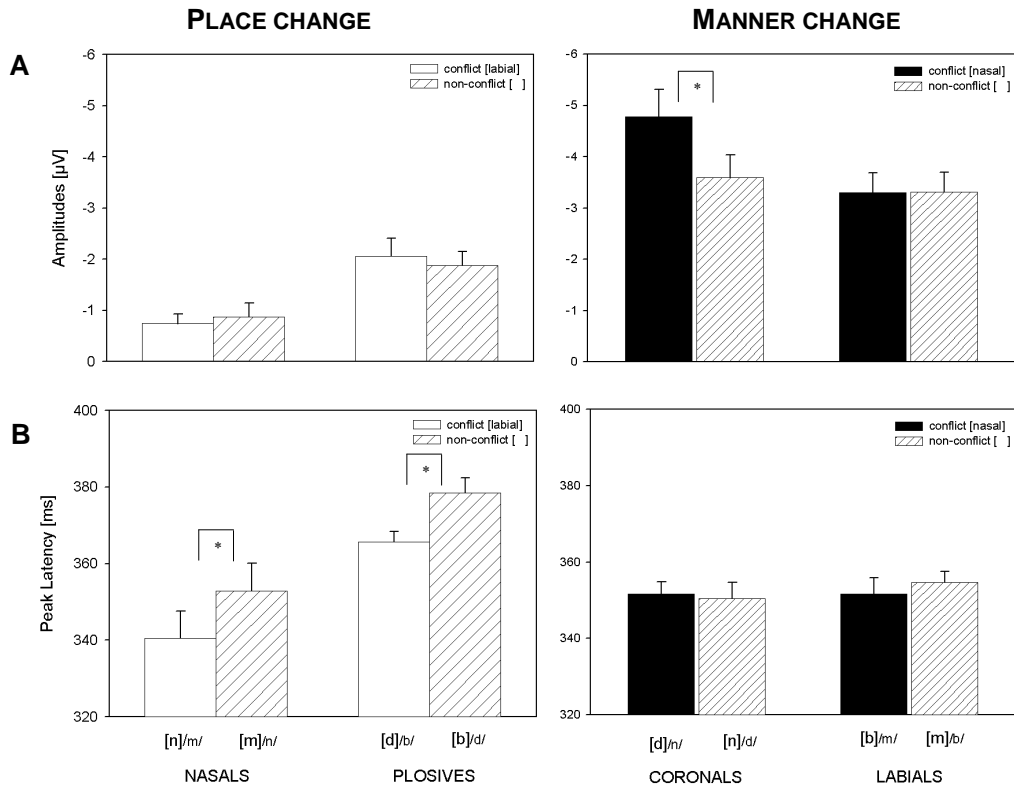


Figure 23: A) MMN mean amplitudes of all eight conditions, the two PLACE CONTRASTS for the nasal & plosive sounds and the two MANNER CONTRASTS for the coronal and labial sounds. B) MMN peak latencies for the eight experimental conditions, the two PLACE CONTRASTS and the two MANNER CONTRASTS. Error markings on top of the bars represent the standard error of the mean (SEM).

P300 Effect

Based on visual inspection of the grand-average waveforms, there was again a large positive deflection at fronto-central electrode sites immediately following the MMN for the coronal manner-of-articulation condition (see Figure 22). This P3-like component was largest at Fz and Cz, yet rather small at Pz, which seems to reflect a P3a (Squires et al., 1975; Knight et al., 1989; Katayama & Polich, 1998). Therefore, further analyses were carried out. As in Experiment 3 and 4 these P3a-like components were parametrized in a similar way as was done for the MMN. The mean amplitude of the positivity at the Fz electrode (re-referenced against linked mastoids) was used as a dependant variable and was measured over a 50 ms time window from positivity onset to offset (400-450 ms). The statistical analyses were restricted to paired comparisons and *t*-test with the four pairs of inversion. Paired *t*-tests (*t*-test, α -level = .05) for the FEATURE-TYPE place change PLOSIVES [LABIAL] ~ [CORONAL] ($[b]_{/d/} \sim [d]_{/b/}$) and place change NASALS [LABIAL] ~ [CORONAL] ($[n]_{/m/} \sim [m]_{/n/}$), as well as the manner change CORONALS [NASAL] ~ [PLOSIVE] ($[n]_{/d/} \sim [d]_{/n/}$) and LABIALS [NASAL] ~ [PLOSIVE] ($[m]_{/b/} \sim [b]_{/m/}$). Amplitude values of the difference waves were again tested against zero in two-tailed *t*-test separately for each condition.

There was only one significant positivity namely for the condition $[d]_{/n/}$ of the CORONAL [NASAL] ~ [PLOSIVE] contrast ($t(23) = 3.58; p < .002$). All other conditions were not significantly different from zero or not positive (all *t*-values < 1; n.s.). This led to the significant difference between the CORONALS [NASAL] ~ [PLOSIVE] ($t(23) = -3.79; p < .001$). The PLOSIVES [LABIAL] ~ [CORONAL] ($[b]_{/d/} \sim [d]_{/b/}$) contrast also differed significantly ($t(23) = 2.78; p < .011$); however, this was based on a significant negativity in this time window of analysis for the condition $[b]_{/d/}$ ($t(23) = -3.30; p < .003$). All other contrasts did not differ significantly (all *t*-values < 1; n.s.). Thus, the only real P3a effect was found when the coronal nasal sound was the standard followed by a coronal plosive deviant (for detailed statistical analysis see Appendices 5.2.2 Statistical Table 18c).

3.2.3.3 Discussion of Experiment 5

Experiment 5 tested the phonological fine structure of mental representations of different speech sounds and different feature dimensions using the mismatch negativity, in order to further examine the functional organization of the mental lexicon. Brain responses to different sounds were compared that differed either in place or manner of articulation, thus comparing different featural dimension and sound classes with each other. In the place contrasts, the comparison focused on the difference between [CORONAL] ~ [LABIAL] using two different sound classes, such as nasals and plosives. Specifically we compared the two nasal sounds $/n/ \sim /m/$, and the two plosive sounds $/d/ \sim /b/$. In our manner contrasts, we

compared the features [NASAL] ~ [PLOSIVE] for coronal and labial sounds. The specific contrasts here were /n/ ~ /d/, both being coronal sounds, and /m/ ~ /b/, both being labial sounds. Note that whenever one feature dimension varied, the other dimension was kept constant. The idea behind the MMN is that the brain responses are assumed to reflect the degree of comparison between the deviant and the standard sound. It is proposed that the standard pre-activates stored mental representations such as phonological representations, whereas the deviant reflects the incoming acoustic surface information. The comparison between the two offers us the possibility to measure the degree of contrast between different sounds and features and gives us insight into how they might be represented. We have proposed an account where mental representations of lexical form are abstract and sparse, where phonological variation and asymmetries caused by phonological assimilation are accounted for through underspecification of lexical form. Not all sounds are assumed to have the same status rather there exist asymmetric relationships between them. The assumption of FUL (Lahiri & Reetz, 2002, 2010) is that mental representations are asymmetric, such that [CORONAL] and [PLOSIVE] are underspecified, the former for place, the latter for manner information, whereas [LABIAL] place and [NASAL] manner of articulation are fully specified. Accordingly, for our two place contrasts the information about [CORONAL] is extracted from the deviant and mapped onto the pre-activated [LABIAL] by the standard, which creates a conflict and should be reflected in a higher MMN activation pattern (or earlier latency pattern) than the reversed non-conflicting condition, where [LABIAL] is extracted from the signal and is mapped onto no information in the underlying representation. Similar asymmetries are expected for our two manner contrasts: when [PLOSIVE] is extracted from /b/ or /d/ it conflicts with the specified [NASAL] of /m/ or /n/ in the underlying representation. In the opposite direction, when [NASAL] is extracted from the deviant and maps onto underspecified information of the standard, non-conflicting conditions arise. The conclusion that conflicting situations should be reflected in larger or earlier MMN responses than non-conflicting situations is based on previous work (Eulitz & Lahiri, 2004; Walter & Haquard, 2004; also Experiment 1, 3 & 4). As already stated, models supporting the assumption that all core representations of lexical form are fully specified and all information of a sound structure is stored (Bybee, 2001; Johnson, 1997; Norris, 1994) would predict equal responses of the MMN or at least differences would occur on the basis of acoustic differences or statistical probabilities. Models that are based on contextual experience, where top-down phonological inference forms an integral part of normal speech perception processes (e.g. Gaskell & Marslen-Wilson, 1996, 1998; Gaskell & Snoeren, 2008), would need to explain MMN differences by contextual properties.

Overall our results are in line with the FUL model and our predictions (see Table 10). We find asymmetric MMN activations for our conflicting ~ non-conflicting place-of-

articulation contrasts for the nasals ($[n]_{/m/} \sim [m]_{/n/}$) and plosives ($[d]_{/b/} \sim [b]_{/d/}$), however, only in the latency measure. The conflicting conditions reflected earlier latencies than the non-conflicting conditions. For our manner-of-articulation contrasts, we found asymmetric activation patterns for the coronal sounds ($[n]_{/d/} \sim [d]_{/n/}$) in the amplitude measure, a higher amplitude for the conflicting condition. There was no difference in amplitude nor latency for the labial sounds ($[b]_{/m/} \sim [m]_{/b/}$).

It is known that basically any discriminable auditory change elicits an MMN (Nätäänen & Picton, 1987; Nätäänen, 2001). We controlled for these factors as best as possible (see methods section 3.2.3.2 *Stimuli*), also by using the identity MMN, where the analysis is based on a comparison of identical stimuli, in their function as standard and deviant (Eulitz & Lahiri, 2004; Pulvermüller & Shtyrov, 2006; Phillips et al., 2000a, 2000b). Furthermore, MMN elicitation tolerates some range of standard-stimulus variation, as long as some pattern is shared by the standards, which is the situation for our stimuli having introduced minimal variance within each class of sounds (Gomes et al., 1995; Huotilainen et al., 1993; Winkler et al., 1990, cited in Nätäänen, 2001; Phillips et al., 2000a).

Can one explain these asymmetries simply on the basis of acoustic characteristics? The two place contrasts used in this experiment comparing the nasal sounds $[n] \sim [m]$ and plosives $[d] \sim [b]$, were acoustically (pitch and intensity) quite similar. Despite the fact that nasals are both sonorants, and plosives are both obstruents, we still find latency asymmetries in our MMN responses. This is very much in line with the study by Walter and Haquard (2004) using the same sound contrasts in their MEG study. They found a significant difference in amplitude and a strong trend for a latency distinction between their acoustically equidistant – regardless of the order in which they were presented – stimulus pairs *ba/da* and *ma/na*. Their asymmetries were similar to ours in that the conflicting conditions led to higher amplitudes or a trend towards being earlier. One difference to our place contrasts, is that we have our critical consonantal contrasts embedded in medial position, whereas Walter and Haquard (2004) had them in an initial position. Furthermore, our results mirror the effects of Experiment 3 of our [CORONAL] \sim [DORSAL] contrast with stops. Again we find differences in rather weak acoustical sounds such as plosives and also nasals however only in latency, not in amplitude.

Indeed the amplitudes of our nasal and plosive contrasts are very small ($< 2 \mu\text{V}$), smaller than the MMNs of the manner contrasts. It might well be that the MMN in fact does not seem to catch the difference as well, especially for the nasals (MMNs below $1 \mu\text{V}$), which could be caused by the way the stimuli were constructed. The perceptual cues for place of articulation in nasal consonants are considered to be provided by formant transitions and nasal murmur. The presence of both cues provides more place information than either

cue presented alone (Kursowski & Blumstein, 1984). In our stimuli the murmur information might be reduced, since they were built in a way that coarticulation effects of the preceding vowel /e/ would be minimized. Therefore, some large part of the transition there was cut off from the nasal consonants, compromising the murmur information. Kursowski and Blumstein (1984) argue ‘that the nasal murmur must be included as a critical acoustic dimension for place of articulation in nasal consonants’ (p. 389). This partly or entirely missing murmur information in our nasals might have led to the fact that the place-of-articulation information between the /m/ and /n/ is not sufficient to elicit larger MMNs in the nasals. For our plosives this would not be the case, and they actually do show significantly larger MMNs than the nasals. The fact that the nasals and plosives both are equal in their acoustic strength plus having no latency differences within their acoustics supports our claim that [CORONAL] and [LABIAL] are differently represented in long-term memory. These findings further support Experiment 3 where we observed asymmetrical activations between our [CORONAL] ~ [DORSAL] stop consonants.

The acoustic characteristics (pitch and intensity; Figure 21) for our manner conditions are quite different, being dissimilar between the two contrasts comparing the coronal sounds [n] ~ [d] and the labial sounds [b] ~ [m]. Here, both contrasts are comparisons between sonorants and obstruents ([n] ~ [d] & [m] ~ [b]), consequently the differences between these contrasts indeed could be based purely on acoustics. However, similar to Experiment 4 the MMN amplitude of the deviant [d] with the standard /n/ is higher than the deviant [n] with the standard /d/, showing the reversed pattern than arguing from an acoustic view. Looking at the other manner contrast where the nasal [m] and the plosive [b] are also acoustically rather different we do not find any differences in the MMNs. This gives even more reason to doubt a purely acoustic reasoning.

Furthermore, since our stimuli had natural timing differences in the onsets between the [PLOSIVE] and [NASAL] sounds (see 3.2.3.2 Stimuli), this could have influenced the MMN latency responses for the two [NASAL] ~ [PLOSIVE] contrasts. However there were no latency differences at all in their MMN response patterns.

So far, one can conclude that the acoustic explanation does not hold in explaining our latency and amplitude differences. The acoustic differences do not seem to be reflected in the manner-of-articulation contrasts, nor do they play a role for the place contrasts, since here the sounds are acoustically very similar.

A further relevant issue affecting the MMN and therefore our results as well might be the influence of phonotactic probabilities. The study by Bonte et al. (2005) argued for auditory cortical tuning to distributional frequencies of phoneme clusters in the language environment.

The distributional probabilities of [VC] sequences in our experiment were highest for the cluster with [b], while all other stimuli were fairly similar (see methods and Table 11). According to the results of Bonte et al. (2005), the prediction for the present study would be a higher MMN amplitude for [ebi] compared to [edi] or [emi]. Or since the probabilities of [eni] are slightly higher than [edi] (only a difference of .56 or .33; see Table 11) one would predict a higher amplitude for [eni]. However, the pattern of our MMN differences was different. The conditions with the largest difference in the phonotactic probabilities showed no MMN differences in amplitudes ($ebi = emi$ and $ebi = edi$), whereas the conditions with a moderate difference showed a reversed MMN difference, larger MMN amplitudes for the low phonotactic probability conditions ($edi > eni$) or no amplitude difference ($eni = emi$). Thus, contrary to Bonte et al. (2005), no MMN amplitude difference was seen between our high vs. low phonotactic probability conditions.

Furthermore, our findings cannot be explained by individual sound frequency effects of our stimuli. The intervocalic frequency turns out to be highest for [n], slightly lower for [d] and lowest for [m] and [b] (see Table 11). Again, a high sound frequency deviant might elicit a higher MMN response; however, in our results the MMN amplitude of the deviant [d] compared to the standard /n/ is increased compared to the reversed condition. Additionally, the lowest sound frequency differences are between [n] and [d] and also between [b] and [m], but here we again find a different MMN pattern: large MMN difference between [n] ~ [d] and equal MMN amplitudes between [b] ~ [m]. Therefore, neither frequency effects, nor phonotactic probability influences can explain the pattern of our results.

Our results are hard to explain in light of surface-based models (e.g. Bybee, 2001), or probabilistic models using contextual and experienced integration (Norris & McQueen, 2008; Gaskell & Snoeren, 2008; Gaskell & Marslen-Wilson, 1996; 1998). The later models especially depend on active processes of compensation for phonological variation through probability computation or recognition of variant stimuli having changed through neighboring contexts, rather than accommodation through representation. In our experiment, we used intervocalic consonantal sounds where the vocalic contexts were purposely left constant. The stimuli were cross-spliced specifically to maintain as much acoustic similarity as possible. Thus, context or experience could not have played a role in the interpretation of these results.

Rather, the results support the notion of underspecification where differences are predicted in the mental representation between sounds. These asymmetries are found in three of our four experimental conditions. Our place contrasts revealed asymmetric activation patterns in the latency range. Our nasal contrast (a) and plosive contrast (b) replicate the MEG study by Walter and Haquard (2004), and is in line with further research using EEG

(Experiment 3; Maiste et al., 1995; Shafer et al., 2004; Walter & Haquard, 2004). These results showed that the MMNs are larger for phonological conflict conditions ([d]_b) as compared to those conditions with no conflict ([b]_d). These asymmetries between [CORONAL] ~ [LABIAL] of the MMN responses can be explained by the underspecification of the place of articulation [CORONAL]. Further evidence comes from Experiment 3 with the same stimuli (VCV-syllables) as in this study comparing [CORONAL] and [DORSAL] place of articulation. We found an enhanced MMN for our conflicting condition [d]_g, when coronal mapped onto the dorsal representation as opposed to the reversed non-conflicting condition [g]_d. The difference to this study was that our [CORONAL] ~ [LABIAL] asymmetries here are reflected in latency asymmetries, not in amplitude asymmetries.

Now coming to our manner-of-articulation contrast: MMN asymmetries were found for the coronal conflict ~ non-conflict contrast (c), but no difference was found for our labial contrast (d). The coronal contrast replicates Experiment 4 finding exactly the same pattern of results, which further supports the assumption of differences in mental representations, such that [PLOSIVE] is underspecified and [NASAL] is specified. The fact that our labial contrast (d) did not show any difference is hard to explain. Nevertheless, acoustic or phonotactic probabilities as influencing factors can pretty much be ruled out. As already discussed above [m] and [b] are quite different in their acoustic characteristics, and phonotactic probabilities, thus having to show a larger MMN differences if one argues as Bonte et al. (2005) has. However this is not the case. It may very well be that the difference could not be captured by the MMN or that labials indeed play an entirely different role. In either case, further investigation is certainly called for.

Another interesting point to discuss is the fact that the MMNs of the place contrasts were again much smaller in amplitude than the MMNs of the manner contrasts. This was also found in the Experiment 3 contrasting place [CORONAL] ~ [DORSAL] and manner of articulation [NASAL] ~ [STRIDENT]. It may be that the manner-of-articulation dimension behaves fundamentally different than the place dimension. In Experiment 4 we claimed that this difference was most likely due to the fact that sounds being altered in the place contrasts are stops and that stops in itself are quite weak sounds. Many studies altering the place of articulation of stop consonants (e.g. Maiste et al., 1995; Shafer et al., 2004; Diesch & Luce, 1997; Pettigrew et al, 2004a, 2004b; Wunderlich & Cone-Wesson, 2001) found very small or no MMNs. At the same time, the nasal contrast also elicited very small MMNs, even smaller than the stops. This indeed could be a consequence of both kinds of sounds being quite weak, but it could also very well be that the place change between the sounds are less perceptive than when the sounds change in their manner of articulation. It remains speculative, but manner of articulation might play a higher-ranking role in the extraction of feature information. Our perceptual system might first extract the information of how a

sound is made, whether it is fully or partially constricted, oral or nasal, before it decides where a sound is articulated, extracting the information about place (Stevens, 2002).

P300 - Automatic Attention Switch

The fifth Experiment revealed the presence of one robust positive component immediately following the MMN, namely the condition where the coronal plosive [d] as deviant followed the coronal nasal [n] as standard. This effect was again more evident at Fz and Cz. It could very well be related to the P3a. The P3a is said to reflect a switch of attention toward the irrelevant deviant when the deviant is considerably different from the context it is presented in (Sams et al., 1985; Escera et al., 2000).

Similar to the previous experiments, this positivity was found for a condition where a sonorant context was disrupted by a non-sonorant deviant. It could be argued that if a context is defined primarily by a sonorant standard, attentional redirection to the non-sonorant deviant would occur by means of frontal lobe activation that may underlie a P3a. In the reversed case, a sonorant deviant and non-sonorant standard discrimination would however diminish this effect because the stimulus change would not demand as much attentional effort so that the deviant sonorant stimulus could more easily be ignored. The coronal plosive deviant seems to mismatch to a greater degree, drawing more attention to the listener as compared to the reversed condition when the deviant is a harmonious sound such as a coronal nasal.

It is suggested, that when the acoustic qualities of sounds differ radically from the qualities of the background stimuli – in this study possibly the difference between the coronal [NASAL] and [PLOSIVE] – the P3 effect reflects the orienting response in the first place with no distinct differentiation of the sound nature per se. It is surprising though that the same contrast with labials did not elicit a P3a in either direction of standard/deviant change. The logic of an automatic attention switch to a non-sonorant sound should be applied here as well. Yet no effect was found.

To sum up, this last MMN Experiment extends the electrophysiological evidence for the concept of specified and underspecified cortical representations for place and manner-of-articulation features. Asymmetric mappings of features extracted from the acoustic speech signal onto their underlying lexical representations for different consonantal sounds and for different feature dimensions have been found. This study sheds light on place- and manner-of-articulation contrasts for consonants in word medial position generalizing previous findings on asymmetrical activations of place features for vowels and consonants (e.g. Eulitz & Lahiri, 2004; Friedrich et al., 2006; Friedrich et al., 2008; Walter & Haquard, 2004; Experiment 1, 3 & 4) such as nasals, plosives and coronals. The resulting asymmetric

activations showed that phonological contrasts at the featural level were predominantly driving the MMN effects rather than other factors of influence such as phonotactic probabilities, sound frequencies, or acoustic factors. Breaking new ground by looking at different feature dimensions with different sounds, this study provides further support for the notion that mental representations of phonological features are not a one-to-one relation between the acoustic speech signal and their mental representations, but are rather more abstract compared to variabilities in the surface realization of lexical forms, as assumed by theories suggesting the storage of all phonological details. These reported MMN asymmetries between conditions for the same acoustic contrasts suggests that the brain refers to underspecified phonological representations as a basic principle in the functional organization of the mental lexicon during speech perception. However, more research is necessary to make further generalizations about this view, for example testing more speech sounds, contrasting more feature dimensions and even contrasting sounds with multiple feature dimensions to test for possible additive and hierarchical effects.

3.2.4 Experiment 6: Reaction Time Experiment

This line of research tested different sound contrasts exploring their lexical representations. The MMN was used as main method in order to tap into these mental representations and to be able to compare the acoustic sound input with the underlying representation. It was further used to compare different mental representations to each other. However, it is quite difficult to distinguish between differences that are phonetic-acoustic and those which are phonological-representational, since they both appear to influence the MMN in amplitude and latency. Sometimes it is not clear whether the MMN is affected by the acoustic differences, which are natural between sounds or whether it indeed is affected by differences in phonological representations, in particular between specified and underspecified phonological representations.

So far the MMN results of the tested sound contrasts are assumed to tap into representational differences finding the predicted fine grained differences in amplitude between [CORONAL] ~ [DORSAL] and [NASAL] ~ [PLOSIVE] sounds, as well as differences in latency between [CORONAL] ~ [LABIAL] sounds. Yet it remains unclear how the acoustic differences between the sound contrasts may have additionally influenced these results. We attempted to control for these acoustic factors between the speech sounds as well as possible, for instance by manipulating the sound stimuli (cross-splicing, loudness, length, etc.) and by using the identity MMN approach and comparing the same stimuli with each other (Pulvermüller & Shtyrov, 2006a; Phillips et al., 2000a; 2000b). Furthermore, some acoustic variability was deliberately introduced to force the perceptual system to map the incoming acoustic signals onto more abstract representations (Eulitz & Lahiri, 2004; Jacobsen et al., 2004; Phillips et al., 2000a, 2000b).

Furthermore, there were also attentional effects that appeared to influence certain sound contrasts. The presence of a robust fronto-centrally distributed positivity immediately following the MMN was seen for those conditions where a non-sonorant sound such as [z], [d], or [g] was the deviant and a sonorant /n/ sound was the standard. These results were consistent across all experiments (Experiments 3-5). Since these positives have a fronto-central distribution and are elicited without a specific task assignment, we are able to identify them as P3a components. The P3a frequently follows the MMN (Lyytinen et al., 1992) and is proposed to reflect an orienting response (Soltani & Knight, 2000; Friedman et al., 2001), an involuntary switch of attention toward unattended changes in the stimulus context when the primary task is not very demanding (Sams et al., 1985a; Escera et al., 2000). Hence, the P3a is thought to reflect an evaluative, conscious, aspect of the orienting response, as opposed to the detection of deviance per se, which appears to be accomplished at a pre-attentive level reflected in the MMN.

If something in the environment catches one's attention, it will be processed in a favored manner and a P3a will be elicited. It can be argued that this also leads to faster reaction times, focusing the attention on a more salient stimulus. There is evidence that speed of response is associated with focused and sustained attention (Salthouse, 1996, 2000; Verhaeghen & De Meersman, 1998). Response speed is the behavioral endpoint of a cascade of neural processes. It has been shown for the visual domain that if a person attends to a location, events occurring at this location are responded to faster, giving rise to enhanced scalp electrical activity (cited in Posner & Petersen, 1990). The ERP typically associated with response speed and stimulus classification is the P300. High amplitudes and short latencies of the P300 were associated with faster responses (Dimoska, Johnstone, & Barry, 2006; Holm, Ranta-aho, Sallinen, Karjalainen, & Müller, 2006).

Additionally, it has been argued that the response time can be related to the MMN latency as well. Tiitinen, May, Reinikainen, and Näätänen (1994) explored the magnitude of frequency deviance, and found that the frequency deviance and reaction time in an active version of the oddball paradigm followed the peak of the difference waveform by a constant delay at all levels of deviation (i.e. the two functions were parallel). These results were used to argue that the MMN reflects the amount of perceived difference between the neural representations of the standard and deviant sound. This is compatible with the hypothesized role of MMN in calling for further processing of deviant sounds (Näätänen, 1990, 1992), since sounds which widely deviate from the preceding sequence are more likely to capture attention eliciting the P3a as correlate of attention switching.

This present behavioral control experiment was aimed at testing acoustic as well as attentional influences on our MMN to all sound contrasts, which have been used in Experiments 3-5. The following discrimination experiment was built to test whether certain sound contrasts differ more acoustically than others since subjects were instructed to react upon sound changes. It was not clear whether the reaction time responses would follow the MMN latency pattern of our previous experiments or rather attentional effects related to the P3a. Therefore the reaction time hypothesis will be laid out for both cases, either showing a similar pattern as the MMN latencies or a pattern related to the P3a.

All sound contrasts such as the two place contrasts [LABIAL] ~ [CORONAL] with nasals ([n]_{/m/} ~ [m]_{/n/}) and plosives ([b]_{/d/} ~ [d]_{/b/}), as well as the place contrast [CORONAL] ~ [DORSAL] with plosives ([d]_{/g/} ~ [g]_{/d/}) were tested in this perceptual discrimination experiment measuring response times. The manner contrasts included the two [NASAL] ~ [PLOSIVE] contrasts with coronals ([n]_{/d/} ~ [d]_{/n/}) and labials ([m]_{/b/} ~ [b]_{/m/}), as well as the coronal [NASAL] ~ [STRIDENT] contrast ([n]_{/z/} ~ [z]_{/n/}). The last contrast included different feature changes depending on the direction of standard/deviant change: a place and a manner contrast [NASAL, CORONAL] ~ [PLOSIVE, DORSAL] ([n]_{/g/} ~ [g]_{/n/}).

Predictions following the MMN latency:

Importantly, all these contrasts were tested in this discrimination experiment to explore their acoustic differences in order to disentangle acoustic-phonetic differences from phonological representational differences. The observed patterns in the MMN results could be reflected in the reaction times of this following behavioral experiment finding faster responses for earlier MMN latencies. If this is the case, it is expected to find equal reaction times for the place contrasts [CORONAL] ~ [DORSAL] ($[g]_{/d/} \sim [d]_{/g/}$), and our manner contrasts [NASAL] ~ [STRIDENT] ($[n]_{/z/}$ vs. $[z]_{/n/}$) and [NASAL] ~ [PLOSIVE] contrast for labials ($[m]_{/b/} \sim [b]_{/m/}$), as these contrasts did not differ in MMN latency. Differences in reaction times are expected for the [NASAL] ~ [PLOSIVE] contrast ($[n]_{/d/} < [d]_{/n/}$) and the combined place and manner contrast ($[n]_{/g/} < [g]_{/n/}$). Following the MMN results, differences in reaction times could also occur between our plosive and nasal place contrasts ($[d]_{/b/} < [b]_{/d/}$ and $[n]_{/m/} < [m]_{/n/}$) as MMN latency differences were found for these contrasts.

Remember that the VCV-stimuli had natural timing differences between the [PLOSIVE] sounds and the [NASAL] and [STRIDENT] sounds. The onsets of [d], [g] and [b] began slightly later than [n], [m] and [z] due to the silent closure duration. This resulted in latency differences between the [NASAL] and [PLOSIVE] contrasts in Experiment 4, showing earlier MMN responses when the [NASAL] was the deviant and [PLOSIVE] the standard. In Experiment 5 these latency differences were no longer present for the two [NASAL] ~ [PLOSIVE] contrasts. For the two place contrasts [LABIAL] ~ [CORONAL] ($[b]_{/d/} \sim [d]_{/b/}$ and $[n]_{/m/} \sim [m]_{/n/}$), they did not vary in their stimulus onsets. The MMN latency effects seemed to be driven by the representational differences due to the underspecified feature [CORONAL]. These latency effects could be mirrored in the response times.

Predictions following attentional effects of the P3a:

It is hypothesized that those contrasts, which had a pronounced P3a-like positivity would be responded to faster as they would attract more attention. Higher attentional arousal would lead to faster response times. Our non-sonorant deviant sounds might attract more attention as they seem more salient than sonorant deviant sounds. Therefore, it is expected that the non-sonorant sounds [z], [d], and [g] are responded to faster when they are presented as deviants in the context of a sonorant /n/ as standard (as they show a pronounced P3a). In order to obtain a more general picture of attentional influences on our sound contrasts, all of them were tested, not only those that showed a pronounced P3a. Consequently, it was assumed that the reaction times between the two place contrasts [LABIAL] ~ [CORONAL] with nasals ($[n]_{/m/} \sim [m]_{/n/}$) and plosives ($[b]_{/d/} \sim [d]_{/b/}$), and the place contrast [CORONAL] ~ [DORSAL] with plosives ($[d]_{/g/} \sim [g]_{/d/}$) should not differ, as they did not elicit positivities following the MMN. The reaction times for the manner contrasts however were expected to

differ since there were pronounced P3a-components found for these contrasts. In particular, the three conditions $[d]_{/n/}$, $[z]_{/n/}$ and $[g]_{/n/}$ are expected to be faster than their reversed conditions $[n]_{/d/}$, $[n]_{/z/}$ and $[n]_{/g/}$ since they seem to cause a switch of attention. The [NASAL] ~ [PLOSIVE] contrast with labials ($[m]_{/b/} \sim [b]_{/m/}$) should not differ in reaction time as there was no pronounced positivity following the MMN.

Table 12: Predictions for the reaction times following the MMN latencies or the P3a amplitudes:

Experimental conditions [Deviant]/Standard/		Predictions following the MMN latency	Predictions following the P3a amplitude
Place Contrasts			
PLACE CHANGE NASALS	$[n]_{/m/}$ $[m]_{/n/}$	$[m]_{/n/} > [n]_{/m/}$	$[m]_{/n/} = [n]_{/m/}$
PLACE CHANGE PLOSIVES	$[d]_{/b/}$ $[b]_{/d/}$	$[b]_{/d/} > [d]_{/b/}$	$[b]_{/d/} = [d]_{/b/}$
PLACE CHANGE PLOSIVES	$[d]_{/g/}$ $[g]_{/d/}$	$[d]_{/g/} = [g]_{/d/}$	$[d]_{/g/} = [g]_{/d/}$
Manner Contrasts			
MANNER CHANGE CORONALS	$[n]_{/d/}$ $[d]_{/n/}$	$[n]_{/d/} < [d]_{/n/}$ (Exp.4) or $[n]_{/d/} = [d]_{/n/}$ (Exp.5)	$[n]_{/d/} < [d]_{/n/}$
MANNER CHANGE CORONALS	$[n]_{/z/}$ $[z]_{/n/}$	$[n]_{/z/} = [z]_{/n/}$	$[n]_{/z/} < [z]_{/n/}$
MANNER CHANGE LABIALS	$[m]_{/b/}$ $[b]_{/m/}$	$[b]_{/m/} = [m]_{/b/}$	$[b]_{/m/} = [m]_{/b/}$
MANNER & PLACE CHANGE	$[n]_{/g/}$ $[g]_{/n/}$	$[n]_{/g/} < [g]_{/n/}$	$[n]_{/g/} < [g]_{/n/}$

All place and manner contrasts used in the previous MMN Experiments (3-5).

3.2.4.1 Methods

3.2.4.1.1 Participants

Altogether 64 right-handed subjects were tested in eight different sound contrasts in this reaction time experiment (age range 19-36; mean age 24.80 (SEM .45); 43 female, 21 male), none of whom participated in the previous MMN experiments. Per sound contrast sixteen subjects were tested.

3.2.4.1.2 *Stimuli*

The experimental stimuli used as standard and deviant in this discrimination task were the same VCV-syllables as in the previous MMN experiments on manner and place features (Experiments 3-5), such as [eni], [emi], [ebi], [edi], [egi], and [ezi] (see Figure 13 & 20 ‘oscillograms’). Pitch and intensity measures were controlled for as far as possible. Figure 14, 17 and 21 illustrates that each set of VCV-syllables show comparable variation in pitch and intensity within the stimuli. The differences between the stimuli are seen in the same figures illustrating the similarities and differences between the contrasts of each set of VCV-syllables.

By again using three variants of each stimulus category some acoustic variability was introduced. Detailed descriptions of these stimuli are provided in the methods section in Experiments 3, 4 and 5 (3.2.1.2 Stimuli of Experiment 3, 3.2.2.2 Stimuli of Experiment 4, & 3.2.3.2 Stimuli of Experiment 5).

In this reaction time experiment, all stimuli were compared in seven sound contrasts with an additional ‘dummy’ contrast, which was used for balancing reasons.²⁴ The seven contrasts were the same as in Experiments 3-5, however, this time in an active discrimination task.

3.2.4.1.3 *Experimental Procedure*

Subjects were either tested alone or two-by-two. Each subject always heard two (of the overall eight) sound contrasts. Sixteen subjects contributed to the average results for each stimulus contrast. They were informed by written instructions that they had to perform a reaction time experiment on sound contrasts presented in a discrimination task. During this task, the subjects were instructed to actively listen to the sound sequences and press a button with their right hand on a response pad as quickly as possible when they detected a deviant sound. The experimental session consisted of four blocks always testing two contrasts. In each experimental block, the nonwords were combined pair-wise, with each serving as a standard and as deviant. During the study, 170 standards (85%) and 70 deviants (15%) were presented for each of the four blocks in one experiment (Table 12).

The discrimination task lasted for about 15 minutes per subject, including breaks. Note that the experimental setup was equal to the MMN studies described in Experiment 3-5. However, due to a generally lower signal-to-noise ratio in behavioral tests, it was possible to use fewer trials per block in this experiment.

²⁴ An additional ‘dummy’ contrast ([z]_{/g/} ~ [g]_{/z/}) was included for an equal experimental design, but was not of further experimental interest. Therefore it was not included in the later statistical analysis.

3.2.4.1.4 *Data Acquisition and Analysis*

Participants were seated in an acoustically shielded room and were instructed to listen carefully to the sound contrasts in order to discriminate the presented deviant stimuli as fast as possible. The auditory stimuli were presented and reaction times were recorded and collected with experimental control software designed for neurobehavioral experiments (Presentation Version 13). Reaction times were collected for the deviants only. Mere responses within a 0-1200 ms interval from the onset of the deviant stimulus were included into the analysis.

For statistical analyses, reaction times to the deviants were used as dependant variables. A univariate repeated-measures ANOVA with the between-subject factor CONTRAST (7 sound pairs) and the within-subject factor CONDITION (2 pairs of inversion e.g. [g]_{/d/} ~ [d]_{/g/}) was calculated. The seven sound contrasts included (see Table 12): The pairs of inversion PLACE CHANGE PLOSIVES [LABIAL] ~ [CORONAL] ([b]_{/d/} ~ [d]_{/b/}) and [CORONAL] ~ [DORSAL] ([g]_{/d/} ~ [d]_{/g/}), PLACE CHANGE NASALS [LABIAL] ~ [CORONAL] ([n]_{/m/} ~ [m]_{/n/}), MANNER CHANGE CORONALS [NASAL] ~ [PLOSIVE] ([n]_{/d/} ~ [d]_{/n/}) and [NASAL] ~ [STRIDENT] ([n]_{/z/} ~ [z]_{/n/}), MANNER CHANGE LABIALS [NASAL] ~ [PLOSIVE] ([m]_{/b/} ~ [b]_{/m/}), as well as the MANNER & PLACE CHANGE [CORONAL, NASAL] ~ [DORSAL, PLOSIVE] ([n]_{/g/} ~ [g]_{/n/}).

3.2.4.2 Results

The overall ANOVA revealed a two-way interaction of CONTRAST x CONDITION ($F(1, 6) = 2.85, p < .013$), as well as a significant main effect for CONTRAST ($F(1, 6) = 17.00, p < .000$). This interaction and main effect allowed for further examination of each sound contrast separately.

All our place contrasts did not differ in reaction times ([b]_{/d/} ~ [d]_{/b/}: $F(1, 15) = 0.11$; [g]_{/d/} ~ [d]_{/g/}: $F(1, 15) = 0.55$; [n]_{/m/} ~ [m]_{/n/}: $F(1, 15) = 2.23$; all $p > .1$; n.s.), which was predicted on the basis of the P3a effects (Figure 24 left panel & Figure 25, left panel).

Our manner contrasts differed in only two contrasts in their reaction times: The two coronal manner contrasts [NASAL] ~ [PLOSIVE] ([n]_{/d/} ~ [d]_{/n/}: $F(1, 15) = 5.70, p < .031$), and [NASAL] ~ [STRIDENT], ([n]_{/z/} ~ [z]_{/n/}: $F(1, 15) = 11.48, p < .004$) differed significantly, showing faster responses when the non-sonorant sound, such as the [d] or [z] was the deviant with /n/ as the standard compared to the reversed condition with [NASAL] as deviant ([d]_{/n/} < [n]_{/d/}; [z]_{/n/} < [n]_{/z/}; (Figure 24, right panel). These results also follow the predictions of the P3a effects (see Figure 25, right panel). The manner contrast MANNER CHANGE LABIALS [NASAL] ~ [PLOSIVE], ([m]_{/b/} ~ [b]_{/m/}: $F(1, 15) = 0.08; p > 0.1$; n.s.), as well as the combined MANNER & PLACE CHANGE [CORONAL, NASAL] ~ [DORSAL, PLOSIVE], ([n]_{/g/} ~ [g]_{/n/}: $F(1, 15)$

= 0.01; $p > 0.1$; n.s.) contrast did not differ in response times (Figure 24, right panel; for detailed statistical analysis see Appendices 5.2.2 Statistical Table 19).

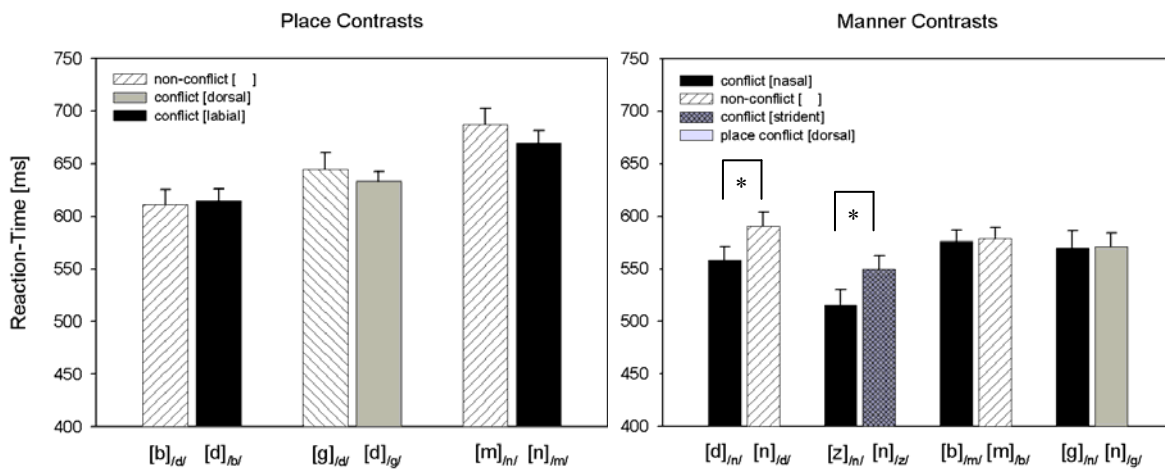


Figure 24: Bar graphs of the mean reaction times for all sound contrasts, the place (left) and manner-of-articulation (right) features. Error markings on top of the bars represent the standard error of the mean (SEM).

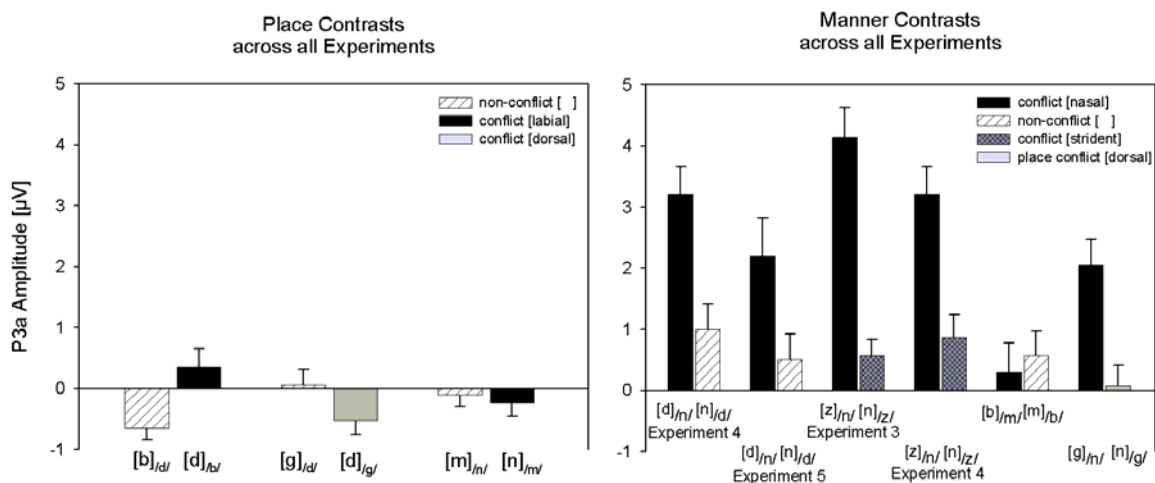


Figure 25: Mean amplitudes of the P3a components for all conditions across Experiments 3-5 at Fz, the place contrasts *Labial-Coronal* [b]_{/d/} ~ [d]_{/b/}, [m]_{/n/} ~ [n]_{/m/}, and *Labial-Dorsal* [g]_{/d/} ~ [d]_{/g/}, as well as the manner contrasts *Nasal ~ Plosive* [d]_{/n/} ~ [n]_{/d/} of Experiment 2 & 3, *Nasal ~ Strident* [z]_{/n/} ~ [n]_{/z/} of Experiment 3 & 4, *Nasal ~ Plosive* [b]_{/m/} ~ [m]_{/b/}, and the combines place and manner contrast *Nasal, Coronal ~ Plosive, Dorsal* [g]_{/n/} ~ [n]_{/g/}. Clearly one can see the pronounced P3a components for the [d], [z], and [g] deviant conditions in the manner-of-articulation contrasts (right panel). The place contrasts do not show any P3a components (left panel), nor does the manner labial contrast.

3.2.4.3 Discussion of Experiment 6

The goal of this behavioral discrimination experiment was to test for acoustic and attentional influences on the MMN of all sound contrasts used in the previous MMN experiments on place- and manner-of-articulation features with consonants. We investigated whether the

reaction time data would parallel latency effects of our MMN, or rather would be explicable on the basis of the attentional P3a effects found in Experiments 3-5. According to the MMN latency predictions, reaction time differences were expected in the two [NASAL] ~ [PLOSIVE] manner contrasts, as well as in the [LABIAL] ~ [CORONAL] place contrasts for plosives and nasals, which was not found. According to the P3a predictions, reaction time differences were only expected in the manner contrasts, particularly in the two [NASAL] ~ [PLOSIVE] contrasts ([d] ~ [n] & [n] ~ [g]), as well as in the [NASAL] ~ [STRIDENT] contrast. Differences were seen for the [n] ~ [d] and [n] ~ [z] contrasts.

These results seemed to have paralleled the results following the P3a amplitude, rather than the MMN latency. In the behavioral experiment, the reaction times seemed to have mapped the P3a attentional effects, in that the two faster response times were those with the highest P3a (see Figure 24 right panel & 25, right panel). The two conditions where the [PLOSIVE] [d] or [STRIDENT] [z] was the deviant with a [NASAL] /n/ as the standard were faster than the reversed conditions. Following this line, the [PLOSIVE] ~ [NASAL] contrast [g]_{/n/} ~ [n]_{/g/} should have shown the same asymmetry in reaction time which was not the case. The other contrasts – in particular the place contrasts – did not differ in their reaction times, nor did they show P3a effects.

The MMN effects of the previous experiments showed amplitude and/or latency differences in certain contrasts, in particular between the place contrasts with plosives ([b]_{/d/} ~ [d]_{/b/} & [g]_{/d/} ~ [d]_{/g/}) and nasals ([n]_{/m/} ~ [m]_{/n/}), but their reaction times did not differ. In a task where subjects were instructed to perceive and react explicitly to acoustic sound differences, no differences were found between the reversals of these place-of-articulation contrasts. This indicates that these sounds are very similar acoustically, which fit to their intensity and pitch contours shown in the methods section (see 3.2.3.1.2 Stimuli & Figure 21). With this result one can very well argue that the MMN indeed taps into representational differences of the different sounds. The differences found in MMN amplitude or latency of the place contrasts can therefore be attributed to differences in representation due to underspecification and not due to acoustic differences as these indeed are minimal.

Regarding MMN effects and reaction times to the manner-of-articulation contrasts, one can argue on the basis of the same logic: The [n] ~ [z] as well as the [n] ~ [d] contrast acoustically differ to a considerable degree (low frequency energy for nasal murmur versus high frequency energy for friction) which is reflected in the reaction time differences, (see Figure 25, left panel). However the MMN for the [NASAL] ~ [STRIDENT] contrast did not differ, and the MMN pattern for the [NASAL] ~ [PLOSIVE] contrast showed mixed results, in that in Experiment 5 the MMN latency did not differ and in Experiment 4 the MMN latency differed minimally but in the opposite direction than the reaction times (RT: [n]_{/d/} > [d]_{/n/}; MMN latency: [n]_{/d/} < [d]_{/n/}). The MMN latency differences of this contrast was originally

attributed to the latency differences in the stimuli onsets (onset differences between [NASAL] and [PLOSIVE]), the reaction time results however show the reversed pattern, supporting stronger attentional influences on this contrast. Overall it seems that the MMN does indeed map representational differences of the sound contrasts pointing to equal representations for the [NASAL] ~ [STRIDENT] contrast and differences in representation between the [NASAL] ~ [PLOSIVE] contrast.

The last two contrasts ([m]_{/b/} ~ [b]_{/m/} and [n]_{/g/} ~ [g]_{/n/}) however are difficult to explain. The MANNER CHANGE [NASAL] ~ [PLOSIVE] with labials does not differ in any measure, neither in reaction time, nor in its MMN pattern, nor in a P3a pattern. The two sounds were expected to be different acoustically and also in their representations since [PLOSIVE] is assumed to be underspecified. This is not apparent in any of the data. The combined MANNER & PLACE CHANGE [NASAL, CORONAL] ~ [PLOSIVE, DORSAL] showed a large MMN difference with a higher amplitude for the conflicting manner condition ([g]_{/n/}) and a large P3a following the MMN compared to the reversed condition. The reaction time results of this contrast do not map this pattern, or the MMN latency difference, which was attributed to the differences in stimulus onsets (onset differences between [NASAL] and [PLOSIVE])

To summarize, the overall reaction time results seem to follow the P3a pattern of involuntary attention shift to sounds that seem acoustically more salient such as [z] and [d] when presented in a harmonious sonorant background such as /n/. When the sound contrasts are acoustically very similar, then the attentional system does not seem to be activated which is shown by the lacking P3as in the place contrasts for plosives and nasals. Additionally, the reaction time results seem independent of the MMN results, in that they reflect only acoustic differences, whereas the MMN on top reflects representational differences. One can therefore assume that the MMN indeed captures the representational level of the speech sounds finding differences between specified and underspecified representations, which are not purely based on acoustic differences. Therefore, this experiment provides further support for the suitability of the MMN method in testing representational characteristics of speech sounds in the mental lexicon. Since the electrophysiological measures do not solely reflect acoustic differences, there is compelling evidence for an abstract underspecified built-up of the sounds under investigation.

4 General Discussion

This doctoral thesis attempted to answer some questions about language perception as for instance, how the constantly varying acoustic speech signal is transformed into a representational format of stored information by a language listener, and how detailed these representations in the mental lexicon might be. Different language perception models have been proposed that deal with these questions diverging in the assumption about the amount and type of information stored. For instance, Bybee argues in her *Usage-based model* for detailed specifications of all phonetic information in stored representations (Bybee, 2001). Individual tokens of experience are stored and organized into categories without removing redundancies. Norris' *Shortlist model* (Norris, 1994) proposes strings of phonemes as primary input that takes the form of phonological representations. These form-based representations are explicitly stored as part of phonological representations in the lexicon. The revised model *Shortlist B* (Norris & McQueen, 2008) computes phoneme probabilities of the incoming signal comparing them against acquired knowledge of phoneme likelihoods in memory. Yet similar to Bybee's model, Norris and colleagues believe that the core representation of a lexical form is fully specified in the mental representation. A third approach assuming fully specified representations is Gaskell's *Distributed Model of Speech Perception* which models the process of speech perception as direct mapping from low-level featural information to distributed representations of lexical knowledge and form. Importantly, this model deals with language variation through phonological inference rules. The context is encountered to infer the underlying stored representation (Gaskell et al., 1995; Gaskell & Marslen-Wilson, 1997, 1996, 1998). Essentially, all models thus far assume fully stored lexical representations, which are based on experience with a language. This assumption does not hold for the *Featurally Underspecified Lexicon Model (FUL)* which is the main focus of this dissertation. This model rather assumes highly abstract lexical representations of features, which are not all stored in the mental lexicon (e.g. Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002; 2010). It is assumed that the acoustic speech stream is converted into distinctive phonological features and directly mapped onto stored lexical representations activating word candidates. All features extracted from the speech signal are compared to their corresponding underlying representations by means of a ternary matching system. The FUL model can account for naturally occurring asymmetries in assimilation processes by assuming underspecified representations combined with its matching logic. The no-mismatch condition in particular allows for variation. Thus, variation in this model is solved at a representational level where abstract features are stored and organized in a hierarchically structured manner. Furthermore FUL is very accurate with

respect to the mapping from acoustic to categorical information, as this mapping is a key process in the perception of naturally varying spoken language.

The assumptions made by FUL of abstract and underspecified lexical representations of different phonological features have been investigated in the work presented here. Many behavioral and electrophysiological studies have already been conducted testing FUL's assumptions regarding mainly place of articulation and height features (e.g. Ghini, 2001; Gumnior et al., 2005; Lahiri & Reetz, 2002, 2010; Wheeldon & Waksler, 2004; Felder, 2006; Felder et al., 2008; Scharinger et al., 2009; Scharinger, 2009; Scharinger & Lahiri, 2010; Eulitz & Lahiri, 2004; Friedrich et al., 2006; Friedrich et al., 2008; Obleser et al., 2003, Walter & Haquard, 2004). Further studies manipulating place of articulation showed asymmetric activation patterns that strongly speak for differences in representation as suggested by FUL (Maiste et al., 1995; Shafer, et al., 2004). Additionally one study by Lahiri and Marslen-Wilson (1991) found asymmetric activations for manner of articulation for vowels cross-linguistically. The experiments conducted here were aimed at generalizing the previously found results in order to test whether the brain indeed refers to underspecified phonological representations as a fundamental principle of the mental lexicon in dealing with language variation and mapping the speech signal onto mental representations for lexical access. More precisely, the goal of these six experiments was to generalize the findings of the underspecified place of articulation feature [CORONAL] and advance this finding to other speech sound classes, and more importantly to generalize specified and underspecified representations to other feature dimensions such as manner of articulation.

Five MMN studies have been conducted with an additional behavioral control experiment to test [LABIAL], [CORONAL], and [DORSAL] place features with different speech sounds such as plosives and nasals, as well as vowels in linguistically more complex structures such as words and nonwords. Crucially, manner features such as [NASAL], [PLOSIVE], and [STRIDENT] embedded in nonwords have been contrasted for the first time in a controlled way.

As a neurophysiological methodology, the Mismatch Negativity (MMN) appeared to be especially apt to capture these assumed representational differences. The MMN as an automatic-change detection brain response reflects – besides acoustic changes – the comparison of surface representations with underlying mental representations activated by the standard (Eulitz & Lahiri, 2004; Ritter et al., 1995). Under this assumption the sound percept that is created by the deviant stimulus corresponds to the surface representation, formed by phonological/phonetic features extracted from the acoustic signal. The series of standards presented before the deviant form a central sound representation that corresponds

to long-term memory traces (e.g. Cowan et al., 1993; Näätänen et al., 1993) and may carry information about the phonological underlying representation in the mental lexicon. Based on this logic the MMN is an excellent tool to study very early automatic sensory processes related to memory representations. Especially the mapping from the acoustic speech signal to lexical representations that seemed to be linked to the auditory cortex can be studied by the MMN (Poeppel et al., 2008) as it is primarily generated in the auditory cortex (Deouell, 2007; Alho, 1995; Näätänen et al., 2007).

Eulitz and Lahiri (2004) already used the MMN successfully to test the nature of mental representations of phonological place of articulation features for isolated vowels. Therefore this study with its results and assumptions made about the MMN and the FUL model provides the basis for this line of research. Eulitz and Lahiri (2004) studied three German vowels in isolation. Following the FUL model, when coronal vowels are presented as deviants, [CORONAL] is extracted from the acoustic speech signal. Presented after a row of dorsal sounds as standards where [DORSAL] is pre-activated in the underlying mental representation a phonological conflict is expected. In the reversed case, then standards and deviants are inverted that is when coronals are standards followed by a dorsal deviant no conflict is expected. In this reversed case of a no-mismatch condition, [DORSAL] is extracted from the signal and mapped onto no pre-activated place information in the underlying representation. Their results showed exactly this pattern: asymmetric MMNs were elicited when they were reversed as standard and deviant in the [CORONAL] ~ [DORSAL] contrast. These asymmetries for similar acoustic differences between pairs of vowels in isolation were interpreted as reflecting the fact that the brain refers to underspecified phonological representations. The results of Eulitz and Lahiri (2004) established the basis for the following experiments.

The *first experiment* in this work attempted to replicate the Eulitz and Lahiri (2004) findings and extend them to linguistically more complex stimuli. For that purpose, the same three vowels were embedded medially in words and nonwords.

The key question was whether vowels in medial positions in complex linguistic structures reliably evoke MMN effects similar to vowels in isolation. Furthermore, it was of interest how possible effects of the lexical status of the stimuli would interact with the phonological level of feature conflicts. The vowel contrasts in words and matched nonwords under study revealed evidence supporting the underspecification of the [CORONAL] place of articulation in vowels. Larger amplitudes were found whenever a place of articulation conflict occurred compared to a non-conflicting situation. This effect was pronounced in the nonword conditions where the lexical effects did not seem to interfere with the phonological level of feature representation. The slightly less pronounced phonological effects for the

word conditions gave us a hint that lexicality may interfere in the relatively early time window of the MMN. As mentioned in the methods section (3.1.1.1.2 Stimuli) our words could not be fully controlled having different meanings or grammatical cases for the three experimental words. These results were interpreted as a possible superimposition of additional factors of influence which could not be fully balanced when using words. Overall the lexicality had no statistical effect, which indicates that there was no difference between words and nonwords. Following these results the choice was made to continue only with nonwords in the following experiments. When looking at the representational level of speech sounds it seemed best to keep potential lexical ambiguities and influences at a minimum.

Importantly, the results of asymmetric MMN patterns for similar acoustic-phonetic vowel contrasts cannot be explained by those models that assume fully specified representations (Bybee, 2001; Gaskell & Marslen-Wilson, 1997, 1996, 1998, Norris, 1994) since they would predict a similar conflict in both directions of acoustic change between the deviant and standard resulting in equal MMN responses.

The reported pattern of MMN differences in Experiment 1 supports the notion that mental representations of phonological place features for vowels are not a one-to-one relation between the acoustic speech signal and their mental representations, but may be more abstract. Theories suggesting the storage of all phonological details cannot account for this.

Experiment 2 was aimed at generalizing these results on place-of-articulation features and the underspecification of [CORONAL] to consonants. Therefore, this experiment contrasted the features [LABIAL], [CORONAL], and [DORSAL] with stop sounds in syllable initial position, with each sound serving as standard and as deviant. Unexpectedly, these acoustically very similar sounds did not show any effects on the MMN. It may very well be that plosives might be acoustically too weak to capture the representational differences. Horvath et al. (2008) argues when the difference between the standard and the deviant sound decreases below the sensitivity of the afferent system, then no MMN can be elicited. At near threshold levels, decreasing MMN amplitudes can be expected.

The lack of pronounced MMNs in this experiment concurs with the findings of other studies examining place of articulation with stops (e.g. Maiste et al., 1995; Shafer et al., 2004; Diesch & Luce, 1997; Pettigrew et al, 2004a, 2004b; Wunderlich & Cone-Wesson, 2001). Place-of-articulation contrasts within plosives seem to be more vulnerable to interference by not only acoustic noise (i.e. they are less discriminable than vowel contrasts) but also by noise caused within the nervous system (Diesch & Luce, 1997; Cowder, 1971). Memory traces that are created by plosives seem weaker than traces created by vowels. Therefore, it is reasonable to suspect that the difference in the effortlessness of discrimination and retention between vowels and plosive stops might affect the amplitude of

the mismatch responses, as it could be the case for the MMN responses in Experiment 2. Furthermore, the speech stimuli were easily discriminable to all subjects when they were attended to, indicating that the poor responses were not caused by qualitatively poor speech stimuli. In particular, the studies by Dalebout and Stack (1999), and Dalebout and Fox (2000) further support the current finding that robust mismatch responses to these fine featural contrasts may not always be obtained although they are easily discriminated and identified behaviorally.

This current study on place-of-articulation contrasts with stop consonants in initial position did not seem to trigger phonological processes or acoustic differences of the F2 formant transitions. At least the information in the initial position did not prove sufficient to generate differences, or to state anything about their representational nature. These plosive sounds presented in syllable initial position without much contextual information may be difficult for the perceptual system to tap into representational differences. The results of Experiment 2 prompted us to try to capture these representational differences in a different context where there is more contextual information available, i.e. with stops in medial position.

In *Experiment 3*, plosive sounds were embedded medially in nonwords contrasting [CORONAL] and [DORSAL]. Since FUL assumes a universal system of phonetic contrasts, not only for place features, but also for other feature dimensions, this third study was aimed at teasing apart effects of phonetic contrasts by systematically manipulating two featural dimensions, place of articulation ([CORONAL] ~ [DORSAL]), and manner of articulation ([NASAL] ~ [STRIDENT]), in order to find further evidence for the generalization of underspecification as a basic principle in the organization of the mental lexicon.

The manner feature contrast was manipulated in order to take a first step into another feature dimension. The question was raised whether manner of articulation features can be studied as successfully as place features using the MMN.

When embedded into vowel contexts, the stop contrasts under study showed evidence for the underspecification of the [CORONAL] place of articulation. Larger MMN amplitudes were found when a place conflict occurred compared to a non-conflicting condition. That is, when the deviant [d] with the extracted feature [CORONAL], mapped onto the pre-activated specified [DORSAL] feature of the standard /g/, a higher amplitude was found than in the reversed case. For the [NASAL] ~ [STRIDENT] manner-of-articulation contrast symmetrical MMNs were found as expected since both are assumed to be fully specified creating a phonological conflict independent of the direction of presentation.

Models that assume fully specified phonological information or models that depend on experience and contextual information (Bybee, 2001; Gaskell & Marslen-Wilson, 1996,

1998) would predict similar conflicts for all directions of change resulting in similar MMN response patterns for the place change contrast.

Experiment 3 provides important results for this work and the following experiments, in that it could show that the MMN is a sensitive measure to test manner-of-articulation feature contrasts. This is the basis for testing fine-grained differences in representation in other feature domains, not only place of articulation. A generalization in other feature areas could now be further explored.

This experiment however also has shown that plosives are unsuitable to test these fine representational differences. Where stops in medial position could still capture these differences between specified and unspecified mental representation despite being very small, stops in initial position could not detect these differences since the MMNs were hardly measurable.

Experiment 4 focused on manner of articulation contrasting the manner features [NASAL], [STRIDENT], [PLOSIVE], in addition to the place features [CORONAL] and [DORSAL] using the same nonwords as in Experiment 3. This study examined the underspecification of the manner feature [PLOSIVE] for the first time. Additionally, an exploratory question was whether manner and place conflicts evoke similar brain responses by testing for differences between manner and place features and whether one particular type of conflict is likely to lead to larger MMNs. The results of this experiment were again inline with the FUL model, where an asymmetric MMN pattern was found for the [NASAL] ~ [PLOSIVE] contrast, and a symmetric pattern for the [NASAL] ~ [STRIDENT] contrast. A higher MMN amplitude was found for the conflicting condition when [PLOSIVE] was the deviant and mapped onto the pre-activated specified [NASAL] in the standard, compared to the non-conflicting reversed case. When both features were conflicting in the case of the [NASAL] ~ [STRIDENT] contrast, equal MMNs were seen, independent of the standard / deviant reversal. Additionally, this demonstrates a replication of the manner contrast of Experiment 3 where the same pattern for the [NASAL] ~ [STRIDENT] contrast was found.

The explorative condition where both manner and place features were varied in the same contrast resulted in an unexpected MMN asymmetry with increased amplitude for the manner conflict than for the place conflict. The plosive deviant with a nasal standard triggered higher MMN amplitudes than the coronal deviant to the dorsal standard. It was speculated that manner of articulation might play a stronger role in the extraction of feature information. First the perceptual system extracts manner of articulation, deciding whether a sound is constricted, completely abrupt, or nasal, and so on, before it extracts information about where a sound is articulated. Linguistic evidence coming from various phonological processes has demonstrated that a more hierarchical organization is necessary to capture properties of sound patterns in natural language (Lahiri & Marslen-Wilson, 1991).

Phonological processes affect subsets of features since they are functionally independent units. Thus, manner features might build a higher class of features than place of articulation features. Lahiri and Marslen-Wilson (1991) mention in their study a system of priorities among different features, such that the manner feature might overwrite the other qualitative features. However, only systematic studies regarding the relationships between different features and feature domains can tell something about hierarchical structures among feature relationships.

Importantly, Experiment 4 illustrates for the first time the underspecification of the manner feature [PLOSIVE] providing a first generalization of the underspecification hypothesis to another feature domain other than place of articulation and height features. Models where all information of sound structure is stored (Bybee, 2001; Johnson, 1997), or models where the canonical forms are fully lexically stored with contextual experience playing a major role in dealing with phonological deviance (Gaskell & Marslen-Wilson, 1996; 1998), MMN differences would not be expected. They could not explain these MMN asymmetries found in the critical [NASAL] ~ [PLOSIVE] contrasts, or the asymmetrical pattern in the combined manner and place contrasts.

The *fifth* and last *MMN experiment* in this work further investigated the fine structure of mental representations for place and manner of articulation features by using the same feature contrasts as in the previous experiments but generalizing them to other speech sounds. The place features [LABIAL] and [CORONAL] were contrasted with stop and nasal sounds. The manner features [PLOSIVE] and [NASAL] were contrasted with coronal and labial sounds. The same nonwords were used as in Experiments 3 and 4.

Overall the results are in line with the FUL model. Asymmetric MMN activations for the conflicting ~ non-conflicting place-of-articulation contrasts for the nasals and plosives were found, however, only in the MMN latency measure. The conflicting conditions reflected earlier latencies than the non-conflicting conditions similar to the study on isolated vowels by Eulitz and Lahiri (2004) and Walter and Haquard (2004) contrasting plosives and nasals in their MEG study. No amplitude differences were found. These two [LABIAL] and [CORONAL] contrasts [b] ~ [d] and [n] ~ [m] are acoustically very similar just as the [CORONAL] ~ [DORSAL] contrast [d] ~ [g] in Experiment 3. Following the assumption of underspecification these three contrasts have differences in their lexical representations, which can be seen in the MMN asymmetries. However it remains unclear why these asymmetries vary in their MMN parameters, showing amplitude differences between the [CORONAL] ~ [DORSAL] contrast and latency differences in the two [LABIAL] and [CORONAL] contrasts. These three contrasts show rather small MMN amplitudes, which might be caused by their acoustic similarities, nevertheless the fine representational differences can be captured in these contrasts for plosives and nasals in medial position.

Another reason for these small MMNs in all place contrasts may be rooted in the way the nasals and plosives are built. As already discussed the perceptual cues for place in nasals (formant transitions and nasal murmur; Kursowski & Blumstein, 1984) for instance might be too weak to capture the place information in the MMN amplitude. Similarly for the plosive sounds, however they actually showed significantly larger MMNs than the nasals. The fact that the nasals and plosives both are equal in their acoustic strength plus having no latency differences within their acoustics (same latency onsets) supports the claim that [CORONAL] is represented differently in the mental representation than [LABIAL] and [DORSAL].

For the [PLOSIVE] ~ [NASAL] manner contrasts, asymmetric activation patterns were found only for the coronal sounds in the amplitude measure. A higher amplitude was elicited for the conflicting condition and a smaller amplitude for the non-conflicting condition. There was no difference in amplitude or latency for the labial sound contrast. The coronal sound contrast provides a replication of the same [PLOSIVE] ~ [NASAL] contrast as in Experiment 4 which strongly speaks for the underspecification of the manner feature [PLOSIVE] however this replication could not be shown for the labial sounds.

Looking at all place and manner contrasts of Experiments 3-5 it is striking that MMN amplitudes of the place contrasts were much smaller than the amplitudes of the manner contrasts. This could be related to the quite weak and acoustically very similar sounds of the place contrasts, but it also very well could be that the place change between the sounds are less perceptible than when the sounds change in their manner-of-articulation domain. It is speculative, but manner of articulation might play a superordinate role in the extraction of feature information as already mentioned above.

Furthermore, manner features differ acoustically more from each other than place features. Also, more features are usually involved in manner differences. These features are also mutually dependent on other features. Manner differences need more featural oppositions.

There is evidence that features are grouped together in a hierarchical structure (e.g. Clements, 1989; Halle & Stevens, 1991). The acoustic-landmark model by Stevens (2002) describes articulator-free features (Stevens, 2002, 2008; Halle, 1992) specifying classes of articulatory actions that are not tied to particular articulators (e.g. vowels, consonants such as sonorant continuant, obstruent, etc.). These induce several types of acoustic landmarks that indicate the presence of segments, and identify regions in the speech signal where acoustic evidence for the articulator-bound features can be found. Articulator-bound features are those features that depend on specific articulators and can be distinctive in a language, such as [ROUND], [HIGH], [LOW], and [BACK] for vowels, and [NASAL], [SPREAD GLOTTIS], etc. for consonants. When the articulator-free features have been established, the numbers of

articulator-bound features fill out the bundles of segments that are needed to identify a sound. Once the cues for an articulator-bound feature have been extracted, these cues must be combined in some way to yield an estimate of the feature. These combinations of cues must be learned by a speaker of a language. The hierarchical extraction process described by Stevens (2002) could be applied to our data where pronounced MMN effects for manner of articulation differences were found. As Stevens proposes a hierarchical features extraction from the acoustic signal, the perceptual system indeed might first focus on a more general classification of acoustic cues based on manner of articulation before deciding on more fine-grained cues relating to place of articulation.

This hierarchical difference between place and manner features might be captured by the MMN. However further experiments and research on this matter needs to be conducted in order to say something about the processing differences between place and manner features.

Note that all the contrasts tested so far have manipulated one type of feature change only. One sound contrast was compared exploring two different feature conflicts within one contrast (Experiment 4; [n] ~ [g]). A place conflict was contrasted with a manner conflict in the reversed condition, and showed an asymmetry in the MMN responses with a stronger response for the manner conflict. Our interpretation of a pronounced influence of manner features compared to place seems plausible, however, further research must be considered to answer questions about hierarchical structures among features. One interesting study by Bonte et al. (2005) found the reversed pattern of results to ours. As already cited in the context of phonotactic probabilities they reported higher MMNs for nonwords with high phonotactic probability (*notsel*) as compared to the low probability condition (*notkel*). Looking closely, they also used place- and manner-of-articulation differences in consonants to create a contrast between standard and deviant, although this was not the aim of their study. At a featural level, the *not[s]el* – *not[k]el* difference included a manner change (from [STRIDENT] to [PLOSIVE]) as well as a place change (from [CORONAL] to [DORSAL]) similar to our *e[n]i* – *e[g]i* difference where we had a manner change (from [NASAL] to [PLOSIVE]) and a place change (from [CORONAL] to [DORSAL]) as well. With respect to the contrasts of place and manner features, the stimuli of Bonte et al. (2005) follow the same systematic as in our experiments. Given underspecification of the [CORONAL] place and the [PLOSIVE] manner of articulation, Bonte's study also has two different feature conflicts in the reversal of sound contrasts. However, they found higher MMN amplitudes for their place conflict (deviant *not[s]el* and standard *not[k]el*). Their reversed manner conflict condition showed reduced MMN amplitudes (deviant *not[k]el* and standard *not[s]el*), contrary to our findings.

It is difficult to conclude anything from these contradicting results. The hierarchical feature structures might also depend on the language of a listener as different languages

might attribute differing relevance to distinct features. Bonte's study was in Dutch and our studies were done with native German speakers. It might be considered that these languages differ in their feature relevance. Either way, place features seem to be fundamentally different from manner-of-articulation features. Certainly, further research is necessary to study these aspects of combined feature conflicts in a controlled fashion. How features of different feature domains are related to each other remains uncertain as this issue has been vitally discussed in the past half century (see Lahiri & Reetz, 2010). However, this type of neurolinguistic research finding correlations of different levels of feature conflicts with the MMN measure certainly opens up new ways to study these feature relations and linguistic structures.

4.1 Possible Influencing Aspects

4.1.1 Latency Effects

It is noteworthy, that Experiments 3-5 demonstrated different results in terms of the MMN parameters, showing effects on the amplitude and/or latency measure. For instance, there were no latency effects in Experiment 3. The representational differences seemed to be mapped onto the differences of the MMN amplitude. Differences in the acoustic onsets between the plosive sounds and the nasal (and strident) did not seem to play a role since the contrasts remained in the same category contrasting only plosives with each other, or nasal with strident sounds.

The representational differences in Experiment 4 also seemed to be reflected in the MMN amplitude. The latency effects which were found for the [PLOSIVE] ~ [NASAL] contrast may reflect the acoustic latency differences between their different stimulus onsets. However this acoustic difference disappeared for the same two [PLOSIVE] ~ [NASAL] contrasts in Experiment 5 where no latency differences were found. Rather the MMN seemed again to map the representational differences by the amplitude difference but only for the coronal contrast in Experiment 5. The latency differences in the place contrasts of Experiment 5 however were attributed to their representational differences since the sounds were acoustically very similar and did not vary in their onset latencies (similar to Eulitz & Lahiri, 2004; Walter & Haquard, 2004). Differences in amplitude were not found.

This rather incoherent picture of latency and amplitude effects between these three experiments seems rather unclear. The representational differences did not appear to be consistently mapped by the amplitude measure, and the acoustic differences between the stimuli were not consistently seen in the latency measure of the MMN. This is rather hard to explain.

The *reaction time experiment (Experiment 6)*, which was conducted as a control experiment does not relate to our MMN latency or amplitude effects. In this behavioral discrimination study, all seven sound contrasts, which have been tested in Experiments 3-5 were compared in an active oddball paradigm. It has been argued that the response time can be related to the MMN latency, in that a frequency deviance and reaction time in an active version of the oddball paradigm follows the peak of the difference waveform by a constant delay at all levels of deviation (Tiitinen et al., 1994). Therefore, this reaction time experiment was aimed at teasing apart acoustic effects as well as attentional influences on the MMN to all consonantal sound contrasts. The question was raised whether certain sound contrasts differ acoustically more than others since subjects were instructed to react on sound changes and that these differences might be mapped onto the MMN latency or amplitude. The reaction times should map the acoustic perceptual difference between the stimulus contrasts.

The results of this experiment did not produce any differences in response time for the place feature contrasts. These results seem plausible since the stimuli of the place contrast (e.g. [n] ~ [m], [b] ~ [d] ~ [g]) indeed are acoustically very similar and close to the perceptual threshold levels. Thus, the place asymmetries found in the MMN amplitudes and latencies can be attributed to the representational differences between specified and underspecified feature representations.

Regarding the manner contrasts, the response time data also show a different pattern than the MMN effects. As for the [NASAL] ~ [STRIDENT] contrast, the deviant [z] was faster than the deviant [n], however, the MMN response did not differ from each other (Experiments 3 & 4). The [NASAL] [n] ~ [PLOSIVE] [d] contrast even resulted in a reversed pattern compared to the MMN (Experiment 4; or Experiment 5 showing no MMN latency difference), showing faster responses for the deviant [d] than the deviant [n]. The combined [NASAL, CORONAL] [n] ~ [PLOSIVE, DORSAL] [g] contrast did not differ in response time which also was not consistent with the found MMN asymmetry (Experiment 4). Hence, similar to the place contrasts, the MMN asymmetries seemed to indeed map the representational differences between the different features, and were not governed by the acoustic differences.

The role of the MMN-generating process suggests that it is involved in adjusting the representation of regularities, the prediction of which was violated by the deviant sound (Winkler & Czigler, 1998; Winkler et al., 1996). The MMN reflects the existence of a representation of stimulus repetition. Violating this regularity does not depend on the amount of acoustic deviation, and therefore the MMN would not only reflect the amount of perceived difference between the standard and deviant sound. For our results, this indeed could mean that the MMN reflects the violation of pre-activated lexical phonological feature

representations, which is reflected in higher MMN amplitudes or earlier latencies by the phonological conflict conditions. In the reversed cases where there is no violation of the lexical representation caused by non-mismatch conditions, the MMNs are smaller in amplitude or later in latency.

4.1.2 *Acoustics*

The MMN is also regarded as a pre-attentive change detection response of the brain, and is considered to reflect the brain's reaction to any change in the acoustic sensory input. We have to ask to what extent have acoustics influenced our results in all of the experiments in this work. The consonantal sounds used in the experiments – [d], [b], [g], [n], [m], and [z] – differ quite substantially in their acoustic properties, particularly in pitch and intensity. One could argue that a strong acoustic cue would trigger a higher MMN response. It has been argued that the MMN amplitude increases and the MMN latency decreases with increased separation between the standard and the deviant, especially when this magnitude of deviance is near threshold levels (Horvath et al., 2008; Tiitinen et al., 1994).

For instance [n] and [m] are sonorant sounds while all others are obstruents, and consequently the differences in MMN pattern between [n] ~ [d], and [n] ~ [g], could indeed be a pure acoustic difference. However, in our results the MMN amplitude of the deviant [d] and [g] compared to the standard /n/ is higher than the deviant [n] compared to standard /d/ or /g/ (Experiments 4 & 5). For an acoustic explanation to hold, one needs to claim that the release of [d] and [g] are acoustically stronger than that of [n], which is not what we see in our stimuli nor in the RT results. There is more reason to doubt a purely acoustic explanation if we look at the [NASAL] ~ [STRIDENT] contrast where the nasal [n] and the strident [z] are acoustically very different (also reflected in the RT data) but their MMN activation patterns are very similar (Experiments 3 & 4). Again, if it were due to the acoustics alone, it would be difficult to argue that the strident and nasal stimuli are equal in their acoustic strength. Furthermore, the consonantal contrasts which are acoustically much more similar to each other such as [d] ~ [b], [d] ~ [g], and [n] ~ [m] show asymmetric activation patterns in the MMN reversals. The RT data strongly support the fact that the found MMN asymmetries are not based on acoustics alone. Especially for the place contrasts it could be shown that they are not perceived differently when they were reversed as standard and deviant, but still showed asymmetric MMNs in either amplitude or latency. Similar arguments against acoustics can be made for the vowel contrasts in Experiment 1. The F2 formant distances of the vowel contrasts [e] ~ [ø] are equidistant to the distances between [o] ~ [ø]. Acoustically this should lead to similar MMN responses within the two vowel contrasts, which was however not the case. The [o] ~ [ø] contrasts elicited an asymmetrical MMN pattern,

whereas the [e] ~ [ø] contrasts were equal which rather speaks for representational than acoustic differences.

Furthermore, with the MMN design calculating the identity MMN and contrasting the sound reversals, it is fairly reasonable to conclude that the acoustic differences between the stimuli are equal and independent of the direction of stimulus presentation. Altogether, there is strong evidence for underspecified mental representations for the features [PLOSIVE] and [CORONAL].

4.1.3 Frequency Effects

Among other possible parameters, which could have affected the MMN results, a possible influence of phonotactic probabilities might be of relevance. Bonte et al. (2005) reported higher MMN amplitudes for nonwords with high phonotactic probability as compared to the low probability condition. As it already has been discussed in the different experiments, the distributional probabilities of [VC] sequences in Experiments 3-5 were lowest for the sequence with [z], while all other stimuli had higher probability distributions ([ez] < [eg] < [em] = [ed] < [en] < [eb]; see probability tables in the different experiments). According to the results of Bonte et al. (2005), the predictions for our experiments would be higher MMN amplitudes for [eni], [edi], [egi], [emi], and [ebi] compared to [ezi]. However, the patterns of MMN differences in these three experiments were entirely different. Additionally, the distributional probabilities of our C[V]C sequences in Experiment 1 were highest for the clusters with [e] and lowest for [ø] stimuli. According to Bonte et al. (2005), the prediction for the vowel study is a higher MMN amplitude for the [e] conditions compared to all the [ø] and [o] conditions. However, the pattern of MMN asymmetries was different in that the conditions with the largest distinction in the phonotactic probabilities showed no MMN differences ([e] ~ [ø]), whereas the conditions with a just moderate distinction in phonotactic probabilities showed a large MMN difference ([o] ~ [ø]).

Furthermore, the findings cannot be explained by individual sound frequency effects of our stimuli. The intervocalic frequency turns out to be highest for [n] and [g], slightly lower for [d], [m], [b] and lowest for [z]. Again, one could argue that a high sound frequency deviant would elicit a higher MMN response; however, in our results the MMN amplitudes show a reversed pattern. The strongest argument against frequency influences is the large divergence in frequency between [n] and [z] without an associated MMN difference at all. Concerning our vowel stimuli in Experiment 1, the individual sound frequencies contrast that [ø] has the lowest frequency, with increasing frequency from [e] to [o]. Following these vowel frequencies, the deviant [o] would have to have the highest amplitude compared to the deviant [ø]. However, the reverse was found in the MMNs. The vowels [e] and [ø] should

also show a difference which was not the case. These results show a pattern, which cannot be explained by frequency effects or influences of phonotactic probability.

Generally, frequency measures must be considered carefully, as many ways are available for frequency to be determined. Distinctions may be related to syllable frequency, co-occurrence frequency, token or item counts, lexical or segment-based frequency. Concerning lexical frequency effects, these can be ruled out altogether, since only nonword stimuli were used in all experiments except for the first study²⁵.

4.1.4 P300 Effects – Automatic Attention Switch

Some of the sound contrasts elicited pronounced positivities immediately following the MMN. These were interpreted as P3a components due to its fronto-central scalp distribution (Squires et al., 1975; Knight et al., 1989; Katayama & Polich, 1998). The P3a component frequently follows the MMN and is thought to reflect the activation of a cerebral network in the orientation of attention toward unattended changes in the stimulus context (Escera et al., 2000; Lyytinen et al., 1992; Sams et al., 1985a). The P3a in our results was elicited whenever a non-sonorant deviant sound such as [d], [g], or [z] followed the sonorant /n/ as standard. The activation of a harmonious sonorant context through the nasal standard could to have been disrupted by the more ‘noisy’ and therefore also more salient non-sonorant deviant, resulting in a switch of attention and an orienting response. The reaction time results seemed to back up this interpretation. The conditions with higher P3a responses also had faster reaction times, as in the case of [d]_{/n/} and [z]_{/n/}. Thus, our response times seemed to have mapped attentional aspects, drawing more involuntary attention to the non-sonorant deviants within the context of sonorant standards. Several language studies have also found the presence of a P3a after the MMN for their sound contrasts. For instance Rivera-Gaxiola et al., (2000) obtained late positive deflections for their native and non-native (Hindi) sounds contrasts that crossed a linguistic boundary. This suggested that that these two conditions could capture involuntary attention plus awareness in the native case. Wunderlich and Cone-Wesson (2001) observed P3as for their deviant and standard speech sounds (words ‘*dæy*’ & ‘*gæy*’ and CVs ‘*dæ*’ & ‘*gæ*’) compared to non-speech tones (tone bursts in the speech frequency range). Pettigrew et al., (2004a) found P3as following their MMN only for certain contrasts ([*gɛy*]_{/gɛ/} & [*gɛ*]_{/gɛy/}). All of these studies suggest that these language sounds engaged the listener’s attention, albeit involuntarily.

Additionally our P3a responses were followed by a reorienting-negativity, the so-called RON component (Schröger & Wolff, 1998a, 1998b), which may reflect redirecting

²⁵ The results of our words in Experiment 1 however did not seem to be governed by frequency effects since words did not differ from our nonwords in the early time window of the MMN.

attention back to the primary task after momentary distraction. The MMN is elicited by incoming sounds that violate some previously invariant characteristics of an auditory stimulus sequence. Thus the MMN is interpreted as a correlate of an automatic, pre-attentive auditory change detection process, while the P3a is interpreted as an index of an attentional shift, and the RON has been proposed to index the subsequent re-orienting of attention to process the task-relevant stimulus information (Beri & Schröger, 2001). This distraction effect is caused by a pre-attentive mismatch detecting process and may be followed by a subsequent attention shift to the changes in the stimuli. In our studies this mismatch detection was followed by a P3a indicating an attention shift to the non-sonorant deviants. The following RON possibly reflected the re-orienting towards task-relevant features.

4.2 Comparison with Predictions of Alternative Models

How do our results fit the predictions made by other models of speech perception? The *Usage-based model* (Bybee, 2001), the *Distributed Model of Speech Perception* (Gaskell et al., 1995; Gaskell & Marslen-Wilson, 1997, 1996, 1998; Gaskell, 2003), and *Shortlist* (Norris, 1994) all assume fully specified mental representations of lexical forms. Bybee argues for detailed stored phonological representations that are based on grouped and categorized units shaped by frequency and experience of language use. Gaskell and his colleagues assume that lexical knowledge is represented in a fully distributed fashion. Variable surface forms of speech are extracted in form of features, which are mapped onto fully specified representations. Norris' earlier version of the Shortlist model assumes that the input consists of strings of phonemes that take the form of phonological representations that are assumed to be explicitly stored. These models have been chosen in comparison to the FUL model since they also focus on the early mapping from the acoustic speech signal onto mental representations in long-term memory.

Furthermore, they all make assumptions about how lexical information is stored in the mental lexicon. These models of speech perception view the core representation of lexical form as being a fully stored representation in the mental lexicon differing in the amount of detail stored. Therefore, these models would predict symmetric activation patterns within the reversals of all sound contrasts. In addition, they all are based on experience with a certain language, ultimately shaped by its frequency of use. Thus, asymmetries occurring in the activation patterns in the reversals of the sound contrasts would then be explained by differences in frequencies. However the models may also have different predictions regarding different frequency measures. For instance Bybee's model relies on frequency of occurrence, but Norris' Shortlist B focuses on phoneme probabilities and likelihood

functions of phonemes and phoneme sequences in a given language. Gaskell's model takes contextual information into account that is, whether a particular sound occurs more frequently in certain contexts. Therefore, depending on the frequency measure, for instance, looking at probability counts, measuring single phonemes, phoneme sequences, phonotactic probabilities, or even looking at frequencies of different phonological features occurring in a language would lead to different asymmetries in the response patterns.

This research considered frequency counts based on token counts of the CELEX database (Baayen et al., 1995) for the critical phonemes and phoneme sequences (e.g. [CV], [C]V, [VCV], V[C]V, C[V]C). The symmetrical and asymmetrical MMN patterns of all five experiments do not seem to be influenced by these measured frequencies. Therefore, context and experience could not have played a role in explaining our MMN results. Rather the results seem to be more strongly influenced by the representational differences as FUL predicts. Table 13 gives a broad overview of all results for the different sound contrasts of all five MMN experiments and the performance of the predictions made by the different models.

Only FUL with its assumption of underspecification of phonological features predicts asymmetries in MMNs within the reversal of sound contrasts, based on differences in representation, not on differences in frequency. These predictions were substantiated by the experiments conducted for this thesis (see Table 13).

Furthermore, Gaskell and Marslen-Wilson's model compensates for variation through contextually sensitive information. It crucially depends on notions of assimilation and recognition of variant stimuli, which have changed due to neighboring contexts. However, our stimuli were deliberately designed in such a way that context remained constant, maintaining as much acoustic similarity as possible.

With respect to more probabilistic connectionist models such as *Shortlist B* (Norris & McQueen, 2008) it seems necessary to take a closer look at whether they might be able to account for our results. Shortlist B also incorporates asymmetries based on prior experience, which are estimated from phoneme-probabilities. These probabilities are compared to the lexical entries that have been estimated from the probability of each word given the path of phoneme-probabilities in the input and the experience with these probabilities. Listeners acquire the knowledge of likelihood functions of phonemes in their language in early childhood. However, this model does not have direct access to the representations of the likelihood functions that listeners have acquired and therefore it is quite difficult to be able to test these representations of likelihood in the way we have done in this work.

Moreover, the model remains vague in actually defining the representations stored in the mental lexicon since its focus lies on the decision making process of selecting the most

probable word, given a certain input over time, as well as the calculations of these probabilities. Thus, Shortlist B is an account for the early phonetic analysis, of how words are matched in the perceptual input and compete with the stored information. Conditional probabilities of each word given the available input are calculated using the Bayes's theorem. The path probabilities can be computed and are compared with each other, which allows for enormous flexibility in the speech signal in dealing with speech variation. The format of representation with which these probabilities are compared against is assumed to be very similar to the original Shortlist version ('A'), where 'the dictionary contains a single canonical representation of each word (i.e. no account of pronunciation variation)' (assumption M3, p. 389; Norris & McQueen, 2008). With this assumption it becomes clear that this model resembles the other models – except for FUL – in that they all take on fully specified lexical representations.

On the other hand, with the new model of Shortlist B one could hypothesize that given the asymmetries in speech, – i.e. asymmetric assimilation processes – the listener has learned these asymmetric probabilities, which may lead to asymmetric representational probabilities in the lexicon. It is to prove whether these asymmetries might be equivalent to the asymmetries found in our MMN results and as FUL would predict them. This actually might be necessary to test all those models that assume that frequency of occurrence influence mental representations in one way or the other (e.g. Bybee, 2001; Gaskell & Marslen-Wilson, 1997, 1996). However as already mentioned, it remains unclear how these representational probability functions can be accessed and calculated in order to compare them against the probabilities of the incoming speech signal.

A further hypothesis might be since Shortlist B relies on computations of probabilities, which are based in experience with a language, one could conclude that certain phonemes or phonological features occurring more often in a language than others would have higher probabilities. According to Paradis & Prunet (1991), coronal sounds for instance are the most frequent sounds in a language. It may be hypothesized that coronals might have the highest phoneme probabilities compared to other sounds. One could argue that a highly frequent sound would trigger a higher MMN response when it serves as a deviant. For our results this would mean higher MMNs for all coronal deviants such as [n], [d], [z], [e], and [ø]. Additionally, our coronal sound contrasts [n] ~ [d], [n] ~ [z], and [e] ~ [ø] should then result in equal MMN responses which was not the case for the [n] ~ [d] contrast, finding higher amplitudes for the deviant [d]. For the mixed coronal and dorsal or labial contrasts [n] ~ [g], [n] ~ [m], [d] ~ [b], [d] ~ [g], and [ø] ~ [o] one would expect higher MMNs whenever the deviant is a coronal sound. This was not consistently the case. For instance, the contrast [n] ~ [g] showed higher amplitudes for the deviant [g] (for an overview see Table 13). For our sound contrasts the results do not necessarily speak for this probabilistic frequency

account, yet it would be worthwhile to further test the exact predictions of each phoneme and phoneme string. For this, the phoneme probabilities and path probabilities of our words and nonwords would have to be calculated and tested. The coronal probabilities for instance also vary and depend on the position in a word and also differ between the sounds (e.g. different frequencies for coronal fricatives vs. coronal stops). More precisely, it would be necessary to calculate the relevant phonemic cues in a certain word position for a particular language, their probabilities and how they are distributed. So far, Shortlist B has only done this for Dutch. The probabilities have not been looked at in other languages. And as already said, how to compare them against the representational probabilities remains uncertain, and would go beyond the scope of this work.

Altogether the format of the stored representations in Shortlist B remains unclear. On the one hand, Norris and McQueen mention them to be very similar to Shortlist A, thus having fully stored mental representations. On the other hand, these representations seem to be learnt in early childhood as probability functions, which might have led to asymmetric representations in long-term memory. Yet, one way or the other our MMN results cannot account for either assumption: Assuming fully stored representations of single canonical representations of each word would lead to symmetric MMN patterns of our sound reversals. Assuming asymmetric representations based on learnt phoneme likelihoods, the MMN asymmetries would have to simulate the probabilistic asymmetries, which are based on frequency measures. Our results do not seem to be influenced by any frequency or probability measure.

So far, only the FUL model with its assumption of underspecification and the ternary-matching logic can account for the MMN asymmetries found within the reversals of sound contrasts in the experiments of this work. Similar MMN asymmetries have been found for place-of-articulation features for isolated vowels (Eulitz & Lahiri, 2004) and consonants (Walter & Haquard, 2004). The MMN results of this research add to a large body of behavioral and neurolinguistic research that has been done on place-of-articulation features finding asymmetric activations. For instance, the underspecification of the feature [CORONAL] has been shown for consonants in medial and initial word position (Friedrich et al., 2006; Friedrich et al., 2008; Walter & Haquard, 2004) in electrophysiological work. Behavioral studies also found asymmetries between [CORONAL] and [DORSAL] segments in medial and final word positions (Lahiri & Reetz, 2002; Wheeldon & Waksler, 2004; Gunnior et al., 2005). This work could show this asymmetry for vowels in linguistically more complex structures (Experiment 1). Furthermore this research could demonstrate asymmetric activations between [CORONAL] ~ [DORSAL] or [CORONAL] ~ [LABIAL] segments in medial position for nonwords for different consonantal sounds, such as plosives and nasals (Experiments 3 & 5). For plosives in syllable initial position this asymmetry was not found

in this work (Experiment 2) for assumed methodological reasons. The really new and groundbreaking findings of this work however are the symmetric and asymmetric activations found for manner-of-articulation features. For the first time it could be shown that the manner feature [PLOSIVE] is a valuable candidate for underspecification similar to the [CORONAL] place feature. When contrasting [PLOSIVE] with [NASAL] asymmetric activations were seen for coronal sounds (Experiment 4 proposed as Cornell et al., *subm.*; Experiment 5). A replication with labial sounds was not observed (Experiment 5) and needs further investigation. When contrasting two specified features such as [NASAL] ~ [STRIDENT], symmetric MMN activations were found despite their large acoustic differences (Experiments 3 & 4).

Table 13: Overview of all experimental contrasts fulfilling the predictions of the FUL model, the Usage-based model, the Distributed-Model of Speech Perception and Shortlist A/B:

MMN Results		FUL (Lahiri & Reetz, 2002; 2010)		Usage-based Model (Bybee, 2001)	Distributed Model of Speech Perception (Gaskell et al., 1995)	Shortlist A/B (Norris, 1994; Norris & McQueen, 2008)
		MMN Amplitude	MMN Latency			
Place contrasts V[C]V-syllables						
PLACE CHANGE NASALS (Experiment 5)	[n] _{nv} < [m] _{n/}	x	✓	x	x	x
PLACE CHANGE PLOSIVES (Experiment 5)	[d] _{b/} < [b] _{d/}	x	✓	x	x	x
PLACE CHANGE PLOSIVES (Experiment 3)	[d] _{g/} > [g] _{d/}	✓	x	x	x	x
Place contrasts [C]V-syllables						
PLACE CHANGE PLOSIVES (Experiment 2)	[d] _{b/} ≥ [b] _{d/}	(✓)	x	(✓)	(✓)	(✓)
PLACE CHANGE PLOSIVES (Experiment 2)	[d] _{g/} = [g] _{d/}	x	x	(✓)	(✓)	(✓)
PLACE CHANGE PLOSIVES (Experiment 2)	[b] _{g/} = [g] _{b/}	(✓)	(✓)	(✓)	(✓)	(✓)
Place contrasts C[V]CVC						
PLACE CHANGE VOWELS	words (Experiment 1)	[e] _{ø/} = [ø] _{e/}	✓	✓	✓	✓
PLACE CHANGE VOWELS	words (Experiment 1)	[o] _{ø/} < [ø] _{o/}	(✓)	x	x	x
PLACE CHANGE VOWELS	nonwords (Experiment 1)	[e] _{ø/} = [ø] _{e/}	✓	✓	✓	✓
PLACE CHANGE VOWELS	nonwords (Experiment 1)	[o] _{ø/} < [ø] _{o/}	✓	x	x	x
Manner contrasts V[C]V-syllables						
MANNER CHANGE CORONALS (Experiments 4 & 5)	[n] _{d/} < [d] _{n/}	✓	x	x	x	x
MANNER CHANGE CORONALS (Experiments 3 & 4)	[n] _{z/} = [z] _{n/}	✓	✓	✓	✓	✓
MANNER CHANGE LABIALS (Experiment 5)	[m] _{b/} = [b] _{m/}	x	x	✓	✓	✓
Manner & Place contrasts V[C]V-syllables						
MANNER & PLACE CHANGE (Experiment 4)	[n] _{g/} < [g] _{n/}	(✓)	(✓)	x	x	x

✓ = predictions fulfilled; (✓) = predictions fulfilled as a tendency; x = predictions not fulfilled; V = vowel; [V] = critical vowel; C = consonant; [C] = critical consonant.

4.3 Questions, Solutions, and Ideas

4.3.1 Open Questions

The MMN experiments reported here have provided compelling evidence for the representational account of FUL, finding asymmetric activations between the reversals of sound contrasts for the underspecified place feature [CORONAL] and manner feature [PLOSIVE]. However some open questions and unresolved issues remain. For instance these asymmetries within the MMN reversals were mainly driven by differences in the MMN amplitude, and sometimes by the MMN latency. Sometimes the latency effects can be attributed to acoustic differences between the stimuli (different onsets of the critical consonants), as in the [n] ~ [d] and [n] ~ [g] contrast in Experiment 4. However these latency differences between the [n] ~ [d] contrast disappeared in Experiment 5. Additionally MMN latency effects arise between the place contrasts, showing latency differences between the [n] ~ [m] and [d] ~ [b] contrast. In these contrasts the acoustics do not differ to the same degree. Here the latency effects are attributed to their representational differences between [CORONAL] and [LABIAL]. Overall it remains unclear why the representational differences sometimes are reflected in the latency measure and sometimes in the amplitude measure. Looking at all results across all experiments, the MMN amplitude, latency and reaction times provide an incoherent picture of the results. Even within similar sound classes the MMN parameters vary (e.g. amplitude difference: [d] ~ [g]; latency difference: [d] ~ [b]). What exactly drives these effects seems uncertain, especially since the other factors influencing the MMN have been well controlled for (e.g. physical parameters of the stimuli such as length, intensity, co-articulation effects, frequency effects, phonotactic probabilities, etc.). Eulitz and Lahiri (2004) could show that the MMN seems sensitive to feature representations in both the MMN amplitude and latency. Walter and Haquard (2004) predominantly found diverging amplitudes corresponding to the differences in feature representations, and a trend towards differences in latency as well. These findings mirror our results in that both the MMN amplitude and latency seem sensitive to representational differences. Yet, what exactly influences which parameter remains unclear and needs further investigation. To actually be able to attribute latency differences to representational differences it seems crucial to control for variation in physical stimulus in future experiments, in particular the onsets of sounds if necessary even through cross-splicing. However then again this will compromise the naturalness of the sounds.

Another unresolved question concerns some ERP effects in Experiment 1. The conflicting word condition [o]_{/o/} elicited a pronounced negativity before the MMN in a time window from 0-200 ms (see Figure 8). This was not seen for the reversed word condition, or

for all other non-conflicting word or nonword conditions. This negativity could be related to a pronounced N100 effect in this condition. The N100 component is associated with the onset of any change in an auditory stimulus, responding to physical stimulus differences. Since this effect is relatively early, starting to differ from its reversed condition already with onset of the stimulus (0 ms), it may be that very early coarticulation effects (which perhaps were not controlled for when building the stimuli) may have played a role in influencing the ERPs. The initial [m] preceding the [o] may have had a different ‘coloring’ than the [m] of the preceding [ø] in that it may be more ‘dorsal’. But then again this effect should have been seen in the nonwords for this condition as well, in particular since the word - nonword stimuli did not differ at this point in time. Hence, this point remains an unresolved issue as well.

One additional matter of concern deals with the phonological feature contrasts that were tested and how they are related to each other in the feature tree proposed by the FUL model (Lahiri & Reetz, 2010). There might be another bias, influencing the manner of articulation contrasts in that the [NASAL] ~ [STRIDENT] and [NASAL] ~ [PLOSIVE] contrasts also oppose sonorant and obstruent. Nasal sounds are sonorants whereas strident and plosive sounds are classified as obstruents. These fundamental different sound classes may further influence our sound contrasts and the MMN responses in a way, which is not clear to at this point. Within the place contrasts, this change is not present since the nasal contrast remains within the sonorants and the plosive sounds remain within the obstruent class. This additional varying factor of sonorant ~ obstruent change between the manner features might influence the MMN in a way that could explain the more pronounced MMNs compared to the much smaller MMNs for the place contrasts. The acoustic and phonological change seems much more pronounced between the manner features, perhaps because they vary between sonorant and obstruents. However, looking closely at this biasing difference and the MMN reversals, this does not seem to influence the MMN responses within the contrasts. As already said, the [NASAL] ~ [STRIDENT] and [NASAL] ~ [PLOSIVE] contrasts all have an additional change of sonorant/obstruent. If this indeed were to influence the MMNs within the sound reversals at a representational level, all four contrasts would have to show the same MMN pattern. This however was not the case: The [NASAL] ~ [STRIDENT] and the labial [NASAL] ~ [PLOSIVE] contrasts elicited symmetrical MMN patterns within the reversals, whereas the [NASAL] ~ [PLOSIVE] contrasts ([n] ~ [d] & [n] ~ [g]) had clear asymmetrical MMN patterns. From this one could conclude that the sonorant ~ obstruent change does not seem to influence the MMN within the reversals at a representational level since the change is the same in either direction of standard/deviant presentation. Yet this change could have influenced the MMN responses in a way that lead to different response

strengths when comparing them with other feature domains such as place of articulation where this change is not present. In order to avoid this biasing super-ordinate sonorant ~ obstruent contrast in future experimental manner contrasts, it might be good to contrast manner features only within the sonorants or only within the obstruents. For instance, contrasting [m] ~ [w] would be a purely sonorant contrast between two specified manner-of-articulation features [NASAL] ~ [CONTINUANT]. To have an asymmetrical contrast it might be of further interest to compare [n] ~ [r], both being sonorants with the same [CORONAL] place of articulation, contrasting the specified [NASAL] with the underspecified [RHOTIC] feature.

4.3.2 *What is Novel about this Research?*

This line of research tested and generalized the specification and underspecification of different features and feature domains as fundamental principle in the functional organization of the mental lexicon. Asymmetric activation patterns between place features such as the specified [DORSAL] or [LABIAL] and the underspecified [CORONAL] sound have been found for different speech sounds and therefore substantiate the assumptions made by the FUL model.

More important are the symmetric and asymmetric activation patterns found for manner of articulation such as [NASAL] ~ [STRIDENT] and [NASAL] ~ [PLOSIVE]. This work provides for the first time electrophysiological evidence for a differential contrast sensitivity of the human brain for physically equalized manner-of-articulation differences during speech perception. Whereas certain manner differences like [NASAL] vs. [STRIDENT] were equally contrastive for both serial orders of presentation, others like [NASAL] vs. [PLOSIVE] turned out to be dependent on the direction of featural change. Among all possible reasons for our MMN asymmetries discussed, the differences in fine structure of phonological representations in the mental lexicon explains the pattern of results best. Taken together, the results support the point of view that the brain refers to abstract phonological representations during speech perception which do not necessarily have to be fully specified, nor are contextually dependant. For the manner of articulation, our results strongly speak for the underspecification of the feature [PLOSIVE].

4.3.3 *Future Outlook, Further Research Questions*

Where do we go from here? This work not only answered some interesting questions but also opened new issues, which are of further interest and can promote further questions and experiments.

For instance the predictions of the FUL model can further be tested not only with nonwords, but also with words. Examining these phonological aspects of speech perception in nonwords is necessary, as they are more controllable. However, words are more natural

units in language processing and the assumptions made by FUL should be tested with neurophysiological methods in more natural language contexts such as words. The featural differences could be tested with minimal pairs as in Experiment 1 with different sounds and features. For instance, place of articulation could be contrasted in three minimal pairs with stops such as [Leder] ~ [Leger] ~ [Leber] contrasting [CORONAL] ~ [DORSAL] ~ [LABIAL]. Manner of articulation could be contrasted initially in the three words [Name] ~ [Same] ~ [Dame], or medially in [Fahne] ~ [Phase] ~ [Pfade] contrasting [NASAL] ~ [STRIDENT] ~ [PLOSIVE]. Also testing minimal pairs in another language than German can be considered, where case and meaning might not vary as much. However it remains difficult to find three minimal pairs with varying sounds and features, which would be necessary for control purposes. Additionally, so many parameters need to be controlled for when working with words (e.g. word frequency, lexico-semantic effects; morphological effects, etc.).

Furthermore, to expand the assumptions of FUL further speech sounds and further features could be tested. For instance, an important control contrast for our [NASAL] ~ [PLOSIVE] and [NASAL] ~ [STRIDENT] contrasts would be to test the MMN responses of the [PLOSIVE] ~ [STRIDENT] contrasts by contrasting [d] ~ [z]. Following FUL's predictions and this line of research (in particular Experiment 4) an asymmetric activation pattern would be assumed, i.e. a higher MMN amplitude or earlier latency for the conflicting case when [PLOSIVE] is the deviant and [STRIDENT] the standard, and a smaller or later MMN for the reversed non-conflicting case.

Additionally other manner features could be tested against each other contrasting the two CONstriction features such as [PLOSIVE] ~ [CONTINUANT], contrasting the labials [b] ~ [w], where one would expect an asymmetric MMN activation pattern, since [CONTINUANT] is assumed to be specified whereas [PLOSIVE] is underspecified. Similarly, one could compare the labials [m] ~ [w] contrasting [NASAL] ~ [CONTINUANT], where one would expect a symmetric MMN activation pattern. This [m] ~ [w] contrasts indeed would be interesting since there would be no additional sonorant ~ obstruent change, only a change in between the features [NASAL] and [CONTINUANT], which presumably are both specified for manner of articulation.

Furthermore the contrasts [PLOSIVE] [d] ~ [LATERAL] [l], [STRIDENT] [z] ~ [LATERAL] [l], and [NASAL] [n] ~ [LATERAL] [l] could be of interest, where only the first contrast would lead to an asymmetric MMN response. The last contrast again would change only within the sonorant class. A further purely sonorant contrast would be the comparison of the specified [NASAL] [n] with the underspecified [RHOTIC] [r] feature where an asymmetric MMN pattern would be expected. The feature [RHOTIC] could be contrasted with [LATERAL] as well.

A related but different issue that may be of interest is the additivity of features and how different features are related to each other, especially across different feature domains. One question to investigate would be whether the MMN is sensitive to multiple feature conflicts. A first grasp on this issue was provided in the [n] ~ [g] contrast in Experiment 4 where two different feature conflicts were present depending on the direction of standard/deviant presentation. Manner-of-articulation conflicts seemed to have a stronger influence on the MMN in our work. This also was seen in the results across Experiments where manner feature contrasts produced much higher MMN responses than place feature contrasts. As already discussed, this very well could be related to the additional sonorant/obstruent change. How these different features and feature domains relate to each other in the sense of hierarchical structures would also be an equation to pursue.

One further control experiment could be conducted which tests similar acoustic differences using non-speech stimuli. These stimuli must be matched to acoustic equivalent properties compared to the linguistic contrasts. For instance one could compare the [NASAL] ~ [STRIDENT] contrast to [low frequency energy] ~ [high frequency energy] contrasts.

Another direction of future research may investigate the exact nature of the translation of the acoustic speech signal into temporally aligned phonetic features (e.g. Stevens, 2008), which are then compared to their corresponding lexical representation in the ternary matching algorithm in FUL. To what degree is this mechanism influenced by segmental probabilities, as it has been proposed by Shortlist B? Do differing languages also differ with respect to what features are extracted from the speech signal?

4.4 Conclusion

This work presented results from several ERP studies that focused on perception and neuronal representation of phonological features of German vowels and consonants. Experiments 1-5 looked at representational differences between place- and manner-of-articulation features by means of the MMN. Experiment 6 reported reaction times to control for various acoustic and attentional confounds in speech sound perception. The crucial findings include:

- Asymmetric MMN activation patterns between the place features [CORONAL] ~ [DORSAL] for vowels in linguistically more complex structures such as nonwords (and words by trend; Experiment 1), and for consonants such as word medial plosives (Experiment 3).
- Asymmetric MMN activation patterns between the place features [CORONAL] ~ [LABIAL] for word medial plosives and nasals (Experiment 5).

- Asymmetric MMN activation patterns between the manner features [NASAL] ~ [PLOSIVE], as well as symmetric MMNs for the [NASAL] ~ [STRIDENT] contrast for coronal consonants (Experiments 3, 4 & 5).
- Experiment 3 was the first study to show a sensitivity of the MMN to controlled manner-of-articulation feature contrasts.
- The MMN experiments on manner-of-articulation features took a first step into another feature domain besides features of the ARTICULATOR node, reporting first evidence for the underspecification of the manner feature [PLOSIVE].
- These MMN asymmetries between acoustic/phonetic equalized sound contrasts demonstrate that the MMN is additionally influenced by the level of conflict between phonological features that are extracted by the deviant with those activated in the mental representation by the standard.

This work gives further neurophysiological evidence that the human speech perception system makes use of *mental representations*, which are abstract in nature and *phonologically underspecified*. This being said, the FUL model is to be preferred over other models of speech perception, since it provides testable predictions for the neurophysiological assessment of speech perception, thereby bridging the gap between theoretical linguistics and cognitive neuroscience of language.

5 Appendices

5.0 Figure and Table Index

- Figure 5.1.1 a: Experiment 1: Place Features (Vowels): Pitch contours
Figure 5.1.1 b: Experiment 1: Place Features (Vowels): Intensity contours
Figure 5.1.2: Experiment 2: Place Features (Plosives): Pitch and intensity contours
Figure 5.1.3a: Experiment 1: Grand-average waveforms of the two conflicting nonword pairs
Figure 5.1.3b: Experiment 1: Grand-average waveforms of the two non-conflicting nonword pairs
Figure 5.1.3c: Experiment 1: Grand-average waveforms of the two conflicting word pairs
Figure 5.2.1d: Experiment 1: Grand-average waveforms of the two non-conflicting word pairs
Figure 5.1.4a: Experiment 2: Grand-average waveforms of the [labial] ~ [coronal] pair
Figure 5.1.4b: Experiment 2: Grand-average waveforms of the [dorsal] ~ [coronal] pair
Figure 5.1.4c: Experiment 2: Grand-average waveforms of the [dorsal] ~ [labial] pair
Figure 5.1.5a: Experiment 3: Grand-average waveforms of the place contrast
Figure 5.1.5b: Experiment 3: Grand-average waveforms of the manner contrast
Figure 5.1.6a: Experiment 4: Grand-average waveforms of the manner contrast [nasal] ~ [plosive]
Figure 5.1.6b: Experiment 4: Grand-average waveforms of the manner contrast [nasal] ~ [strident]
Figure 5.1.6c: Experiment 4: Grand-average waveforms of the manner & place contrast
Figure 5.1.7a: Experiment 5: Grand-average waveforms of the place contrast with plosives
Figure 5.1.7b: Experiment 5: Grand-average waveforms of the place contrast with nasals
Figure 5.1.7c: Experiment 5: Grand-average waveforms of the manner contrast with coronals
Figure 5.1.7d: Experiment 5: Grand-average waveforms of the manner contrast with labials
Figure 5.2.1: Grand-average waveforms of the standards for words & nonwords in the conflicting & non-conflicting conditions

5.2.2 Statistical Tables

- Table 14a Experiment 1: MMN Mean Amplitudes
Table 14b Experiment 1: MMN Peak Amplitudes
Table 14c Experiment 1: MMN Peak Latencies
Table 14d Experiment 1: Standard & Deviant Analysis for the N400 Time Window
Table 15a Experiment 2: Individual MMN Mean Amplitudes of the Early MMN & Late MMN
Table 15b Experiment 2: Condition-Related MMN Mean Amplitudes (Early MMN at Fz & Cz)
Table 15c Experiment 2: Condition-Related MMN Mean Amplitudes (Late MMN at Fz & Cz)
Table 15d Experiment 2: Individual MMN Peak Latencies of the Early & Late MMN
Table 16a Experiment 3: Individual MMN Mean Amplitudes at Fz, & Condition-Related MMN Mean Amplitudes at Fz & Cz

Table 16b	Experiment 3: Individual MMN Peak Latencies
Table 16c	Experiment 3: P3a Grand-Average Time Window
Table 17a	Experiment 4: Individual MMN Mean Amplitudes at Fz, & Condition-Related MMN Mean Amplitudes at Fz & Cz.
Table 17b	Experiment 4: Individual MMN Peak Latencies
Table 17c	Experiment 4: P3a Grand-Average Time Window
Table 18a	Experiment 5: Individual MMN Mean Amplitudes at Fz, & Condition-Related MMN Mean Amplitudes at Fz & Cz.
Table 18b	Experiment 5: Individual MMN Peak Latencies
Table 18c	Experiment 5: P3a Grand-Average Time Window
Table 19	Experiment 6: Statistical Analysis of the Reaction Times

5.1 Figures

5.1.1 Experiment 1: Place Features (Vowels): Pitch & Intensity Contours

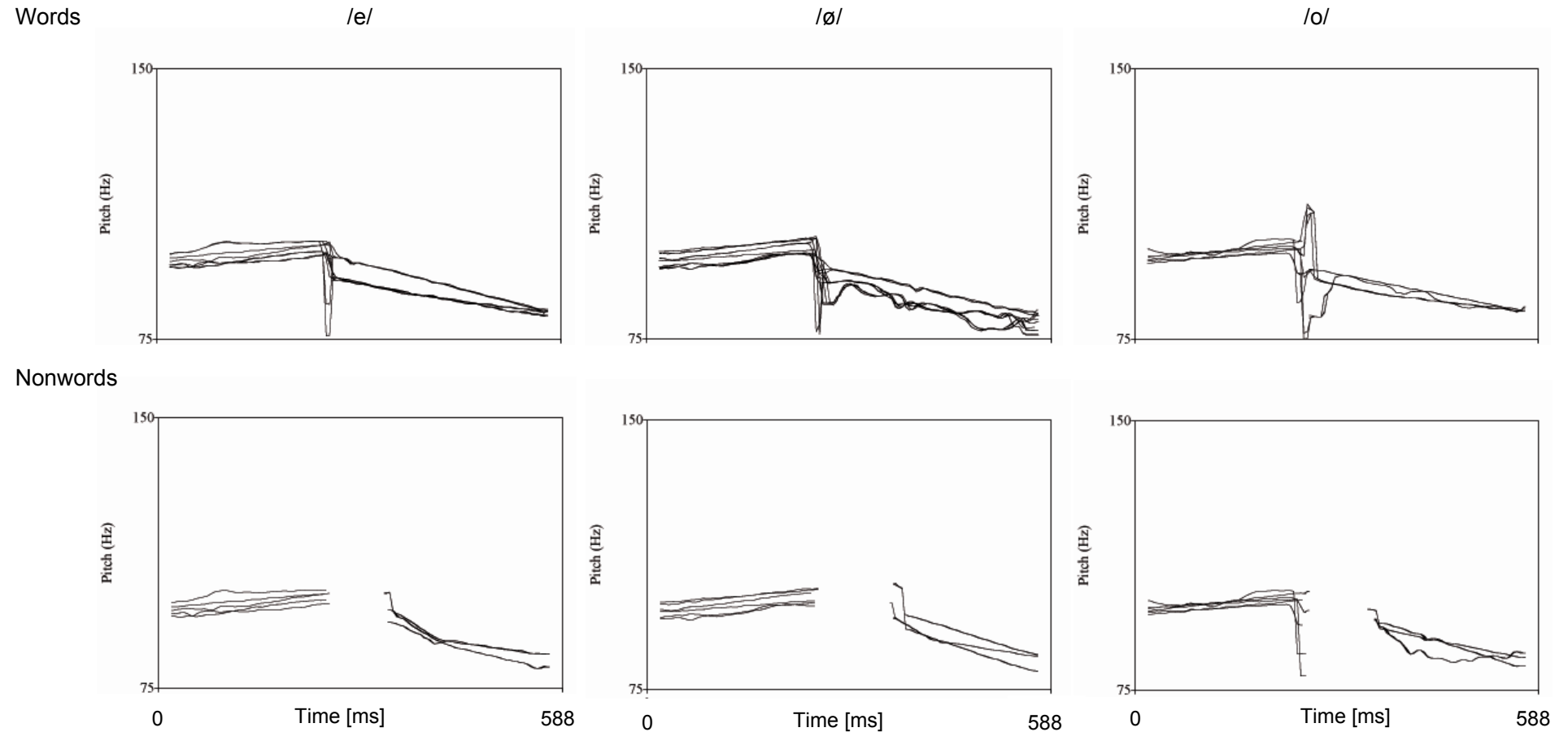


Figure 5.1.1a: Pitch contours of the words [mø:rən], [me:rən], [mø:rən] and the nonwords [mø:kən], [me:kən], [mø:kən] used in Experiment 1. Six variants for each word and nonword were used. Note that the [m] sounds preceding each vowel are the same for words & nonwords. Each set of stimuli showed comparable variation in pitch.

5.1.1 Experiment 1: Place Features (Vowels): Pitch & Intensity Contours

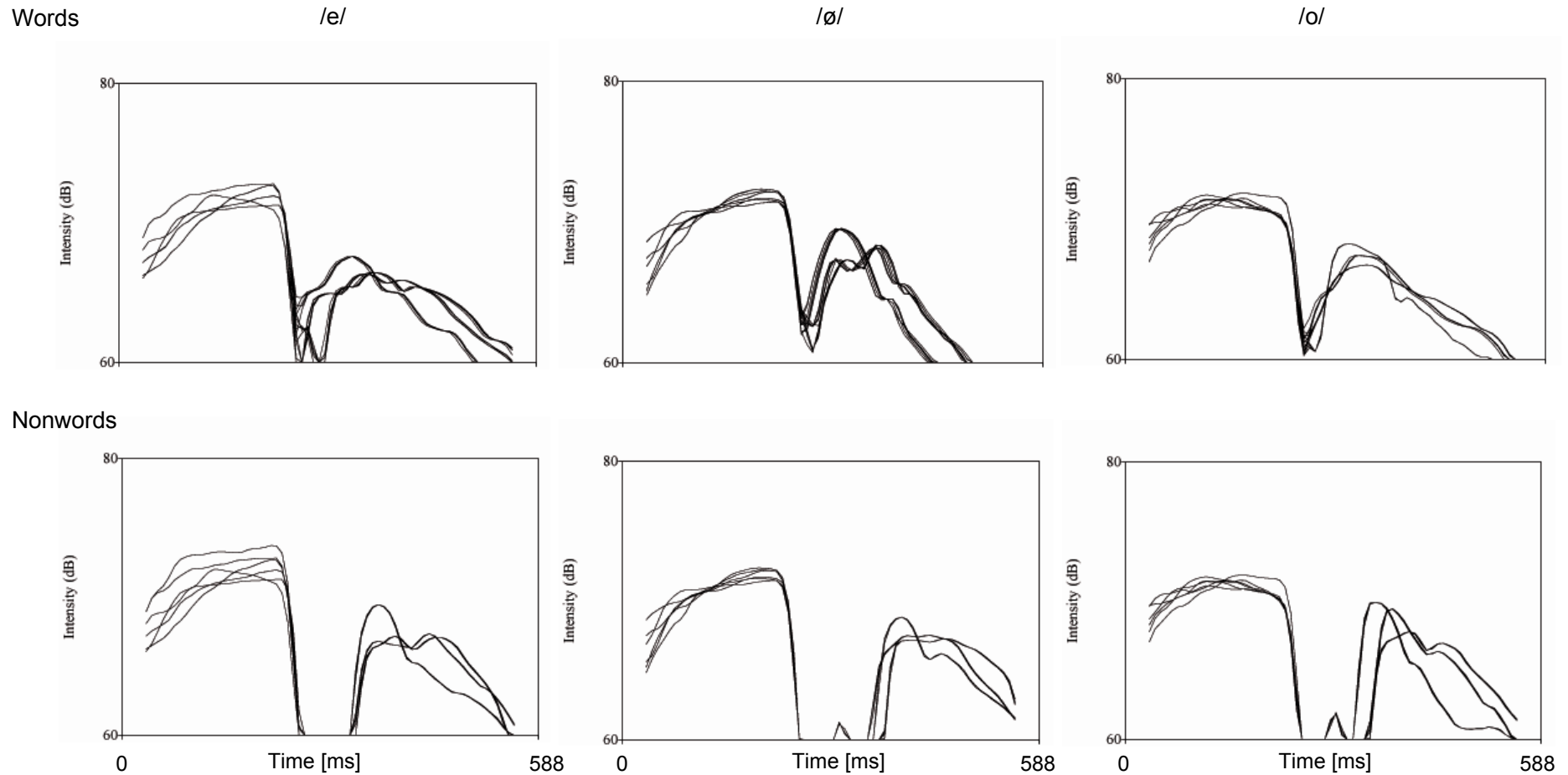


Figure 5.1.1b: Intensity contours of the words [mø:rən], [mɛ:rən], [mɔ:rən] and the nonwords [mø:kən], [mɛ:kən], [mɔ:kən] used in Experiment 1. Six variants for each word and nonword were used. Note that the [m] sounds preceding each vowel are the same for words & nonwords. Each set of stimuli showed comparable variation in intensity.

5.1.2 Experiment 2: Place Features (Plosives): Pitch & Intensity Contours

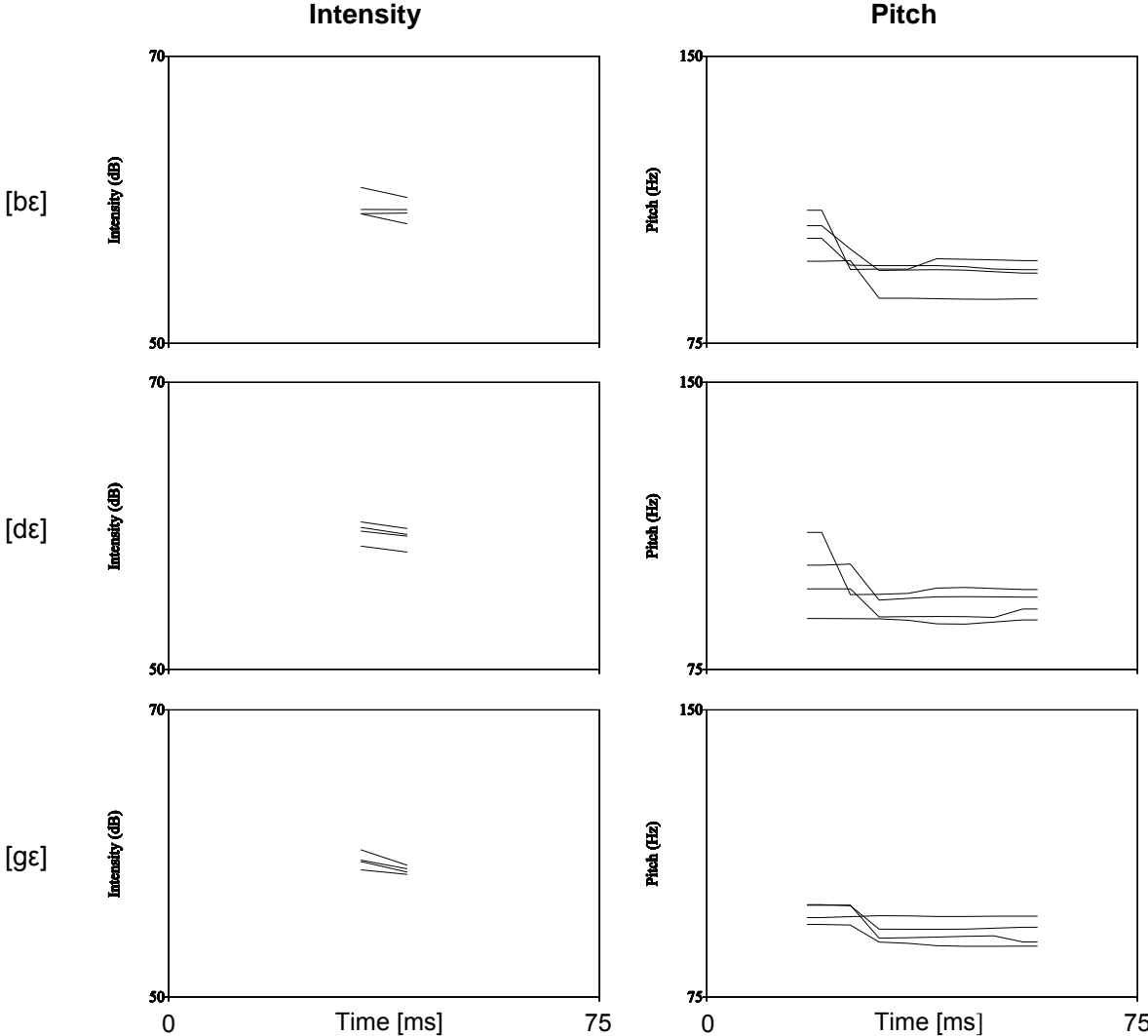


Figure 5.1.2: Pitch and intensity contours of the CV-syllables used in Experiment 2: [bɛ], [dɛ], and [gɛ]. Four variants for each CV-syllable were used, each set of syllable showed comparable variation in pitch and intensity.

5.1.3 Experiment 1: Place Features (Vowels): Standards, Deviants, & MMN Waveforms

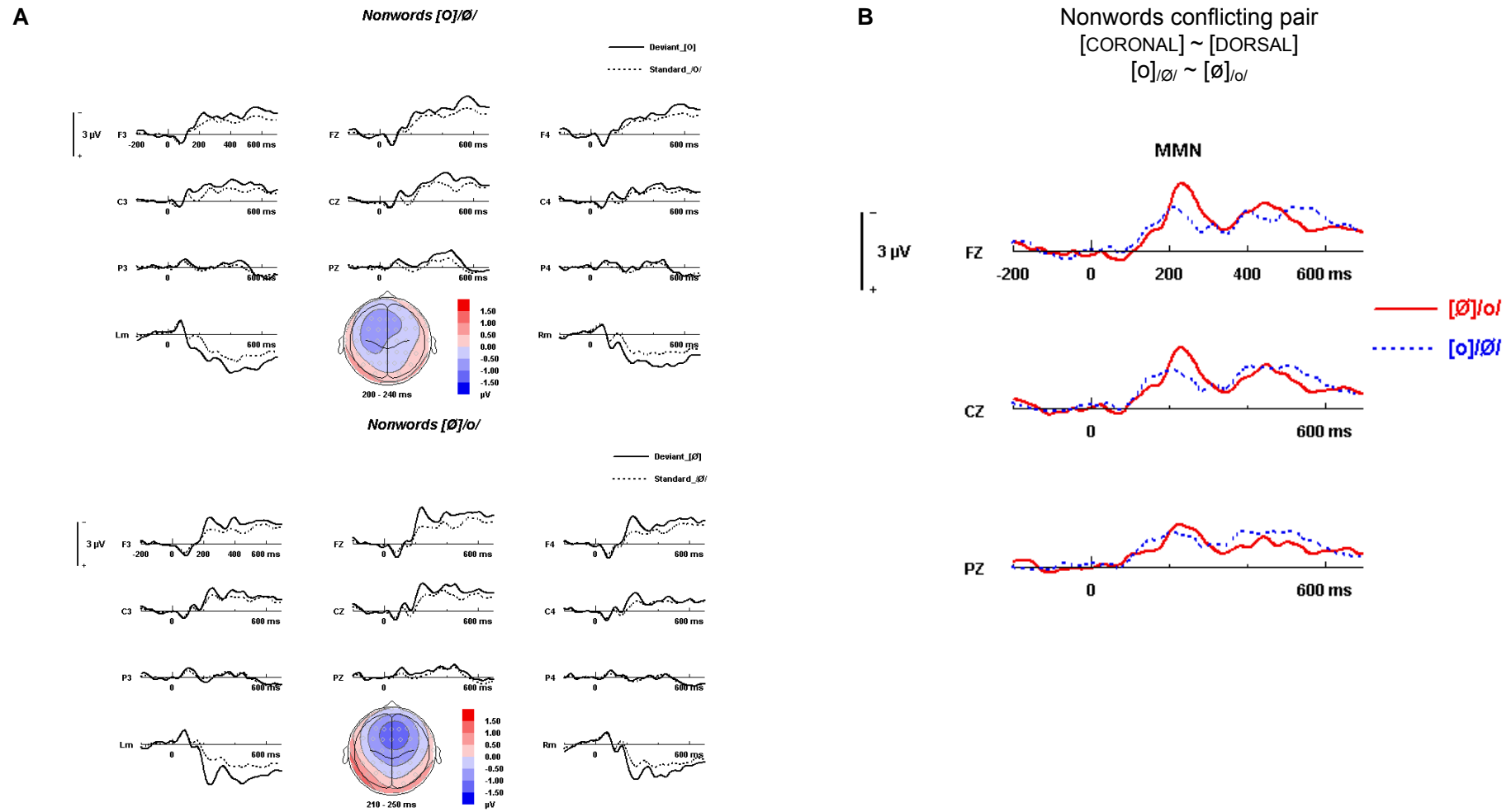


Figure 5.1.3a: Grand-average waveforms for the two conflicting nonword pairs, [o]_∅/ and [∅]_o/. (A) shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). (B) shows the difference wave of the conflicting vowel contrasts for nonwords at the midline electrodes (Fz, Cz, Pz).

5.1.3 Experiment 1: Place Features (Vowels): Standards, Deviants, & MMN Waveforms

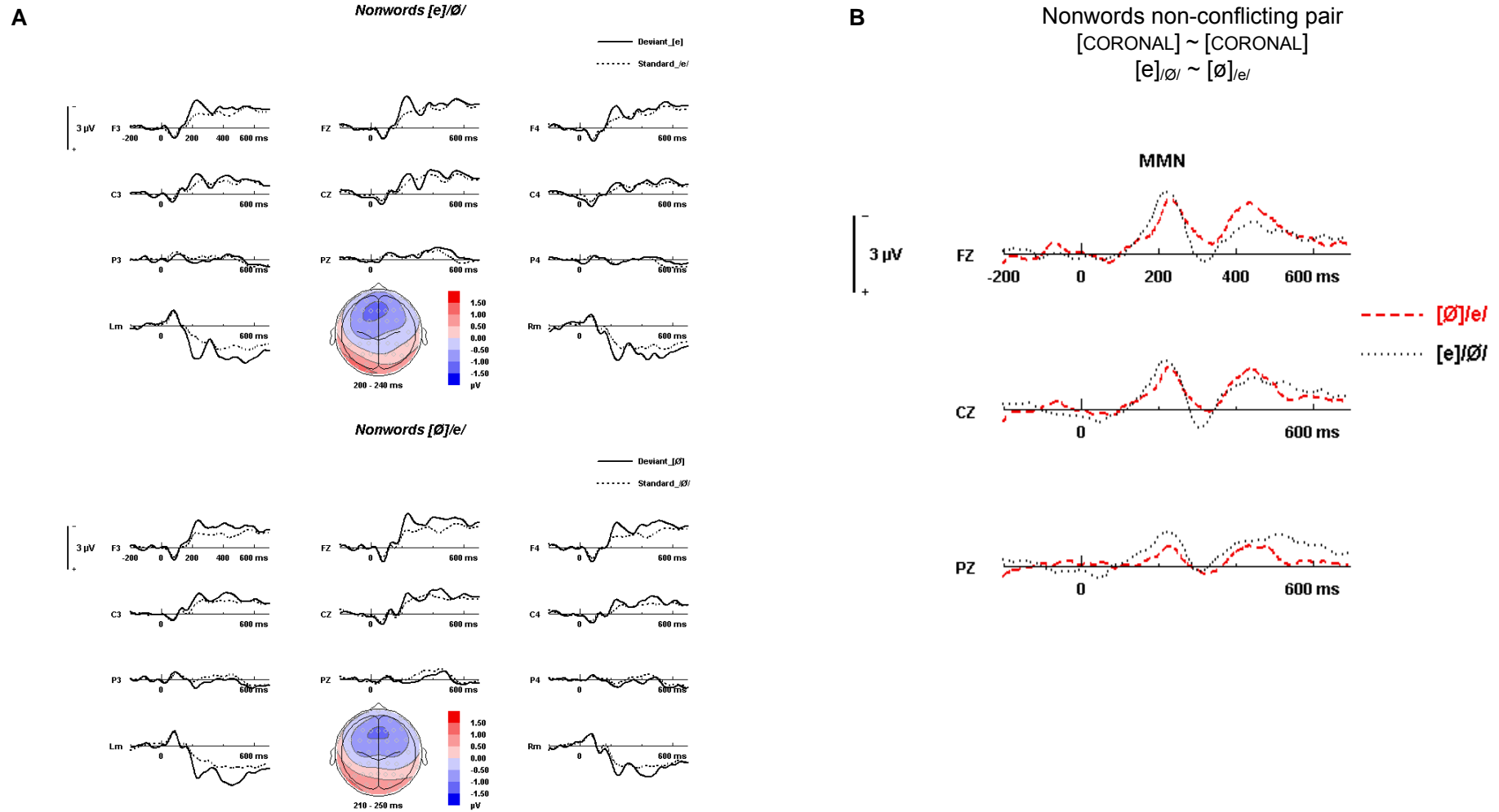


Figure 5.1.3b: Grand-average waveforms for the two non-conflicting nonword pairs, [e]_/[Ø] and [Ø]_/[e]. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the non-conflicting vowel contrasts for nonwords at the midline electrodes (Fz, Cz, Pz).

5.1.3 Experiment 1: Place Features (Vowels): Standards, Deviants, & MMN Waveforms

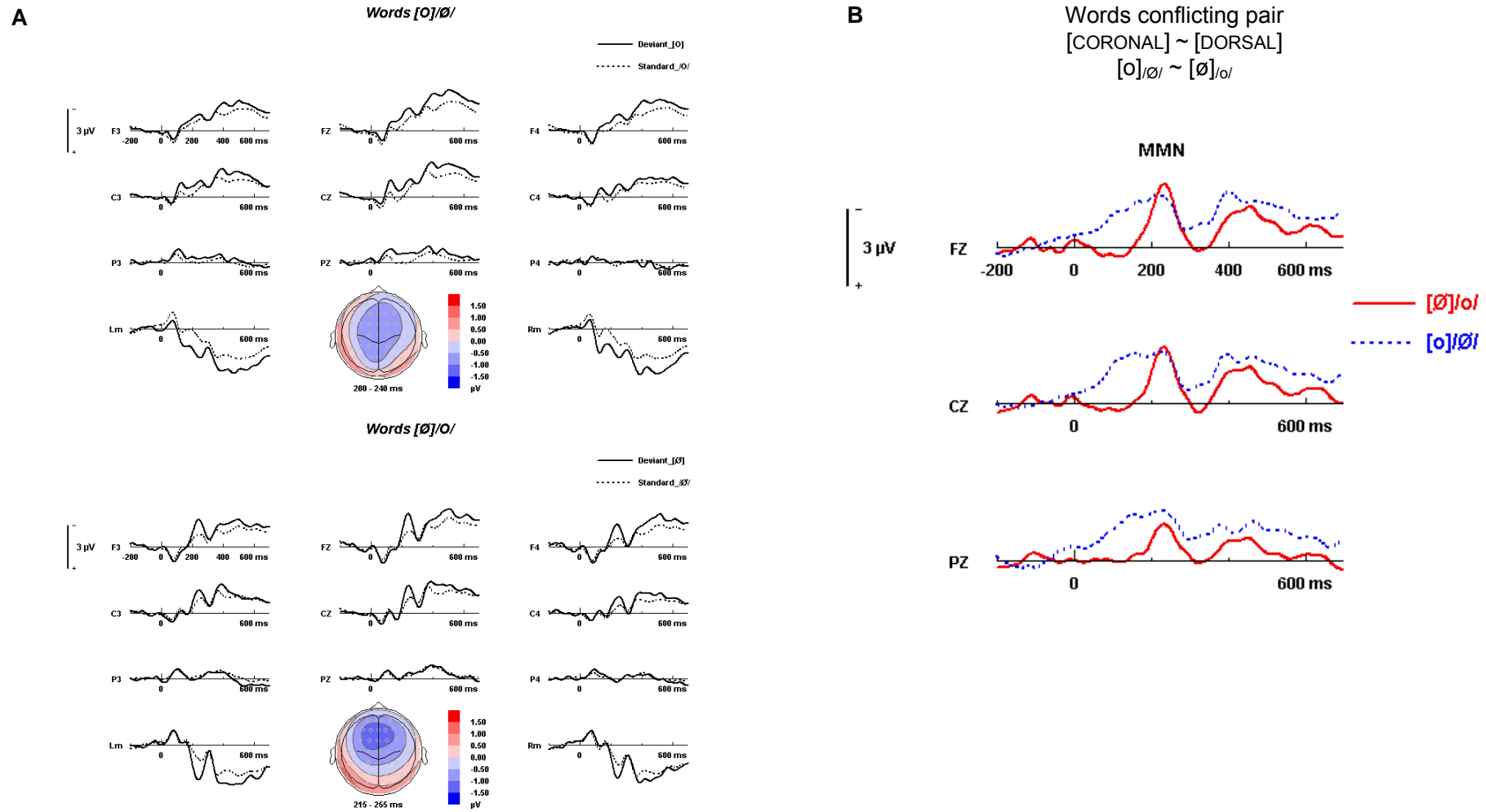


Figure 5.1.3c: Grand-average waveforms for the two conflicting word pairs, [o]_{o/} and [ø]_{ø/}. (A) shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). (B) shows the difference wave of the conflicting vowel contrasts for words at the midline electrodes (Fz, Cz, Pz).

5.1.3 Experiment 1: Place Features (Vowels): Standards, Deviants, & MMN Waveforms

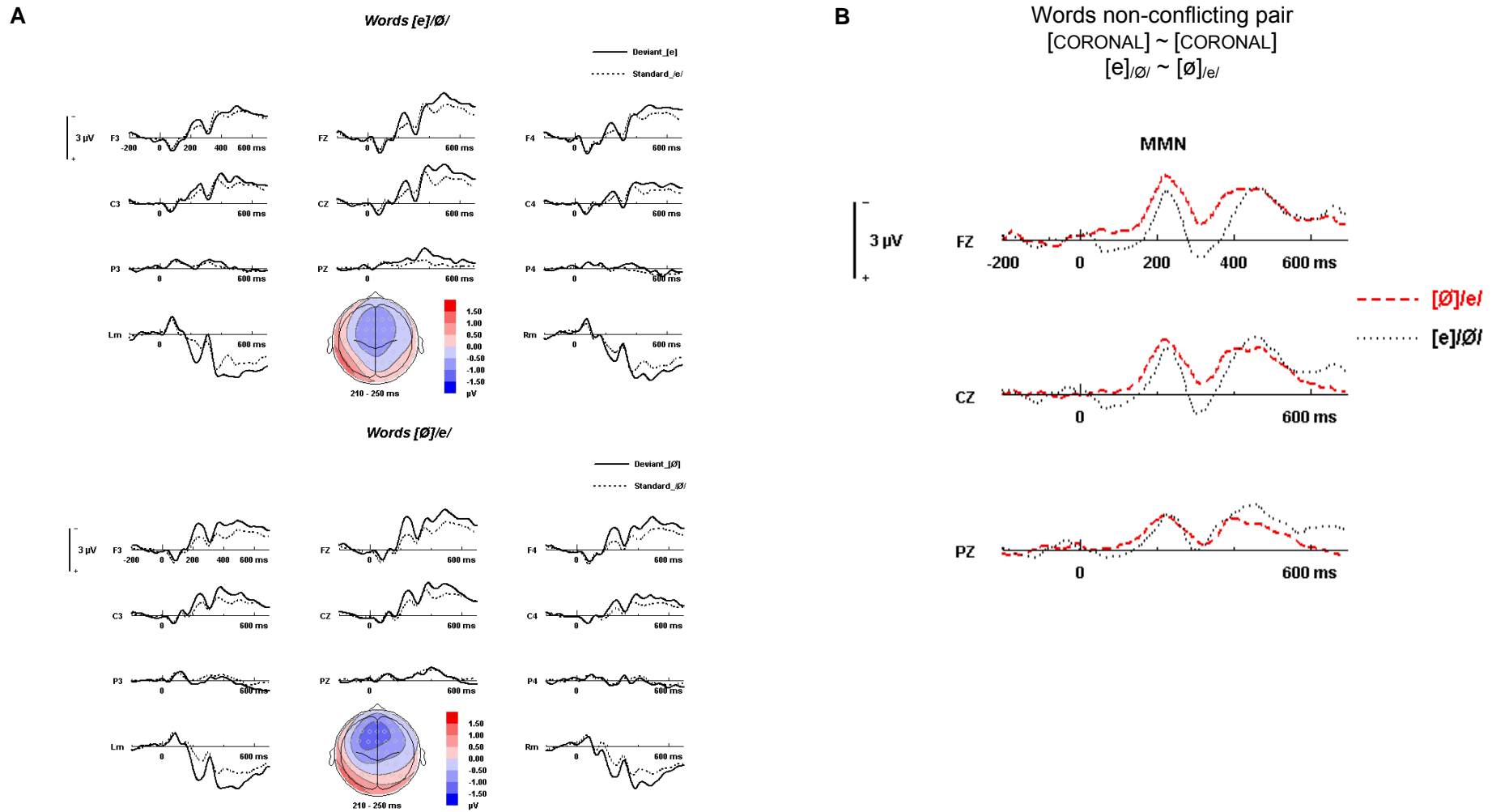


Figure 5.2.3d: Grand-average waveforms for the two non-conflicting word pairs, [e]_{/Ø/} and [Ø]_{/e/}. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the non-conflicting vowel contrasts for words at the midline electrodes (Fz, Cz, Pz).

5.1.4 Experiment 2: Place Features (Plosives): Standards, Deviants, & MMN Waveforms

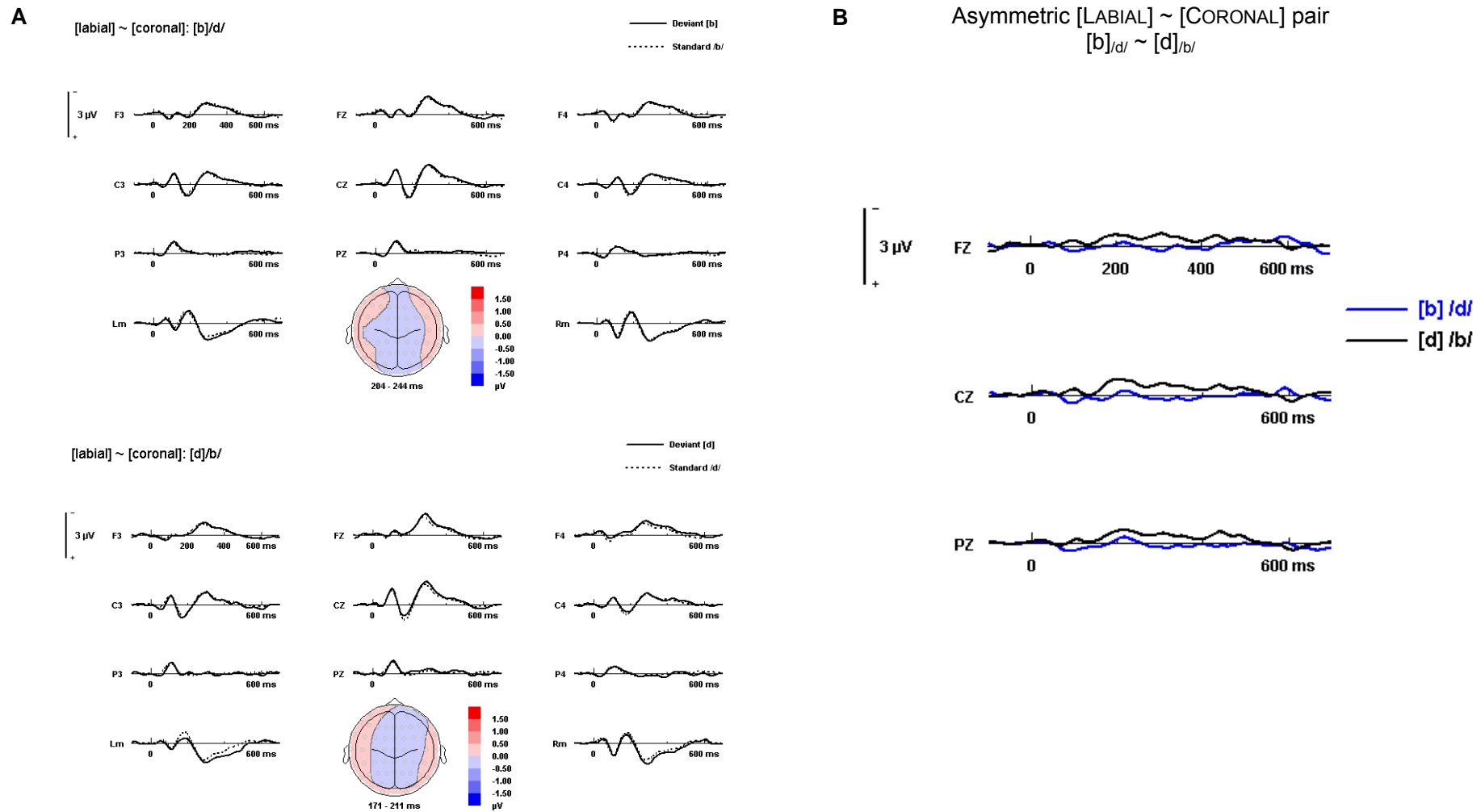


Figure 5.1.4a: Grand-average waveforms for the [LABIAL] ~ [CORONAL] pair, [b]_{/d/} and [d]_{/b/}. (A) shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). (B) shows the difference wave of the asymmetrical [LABIAL] ~ [CORONAL] contrast for plosives at the midline electrodes (Fz, Cz, Pz).

5.1.4 Experiment 2: Place Features (Plosives): Standards, Deviants, & MMN Waveforms

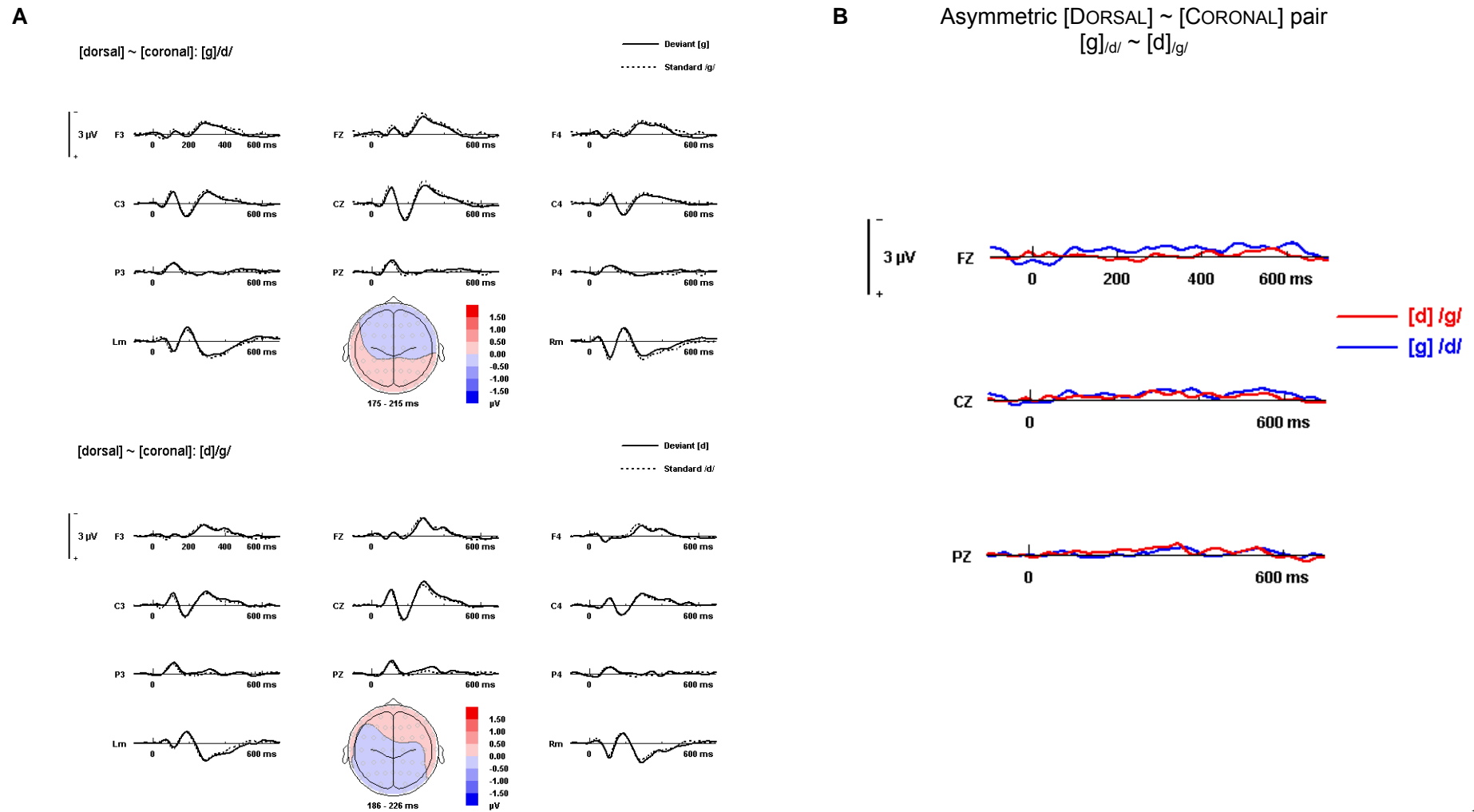


Figure 5.1.4b: Grand-average waveforms for the [DORSAL] ~ [CORONAL] pair, [g]/d/ and [d]/g/. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the asymmetrical [DORSAL] ~ [CORONAL] contrast for plosives at the midline electrodes (Fz, Cz, Pz).

5.1.4 Experiment 2: Place Features (Plosives): Standards, Deviants, & MMN Waveforms

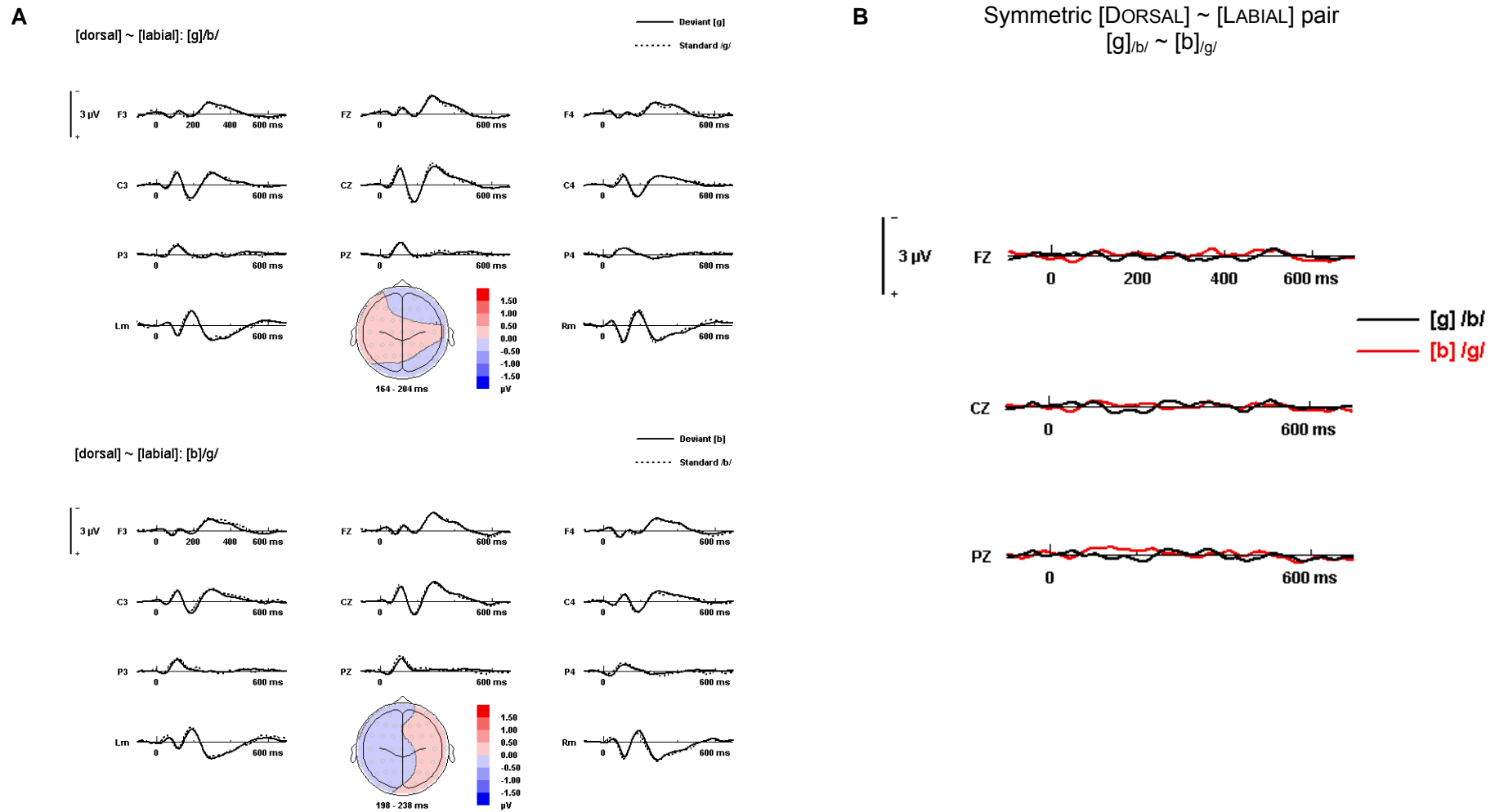


Figure 5.1.4c: Grand-average waveforms for the [DORSAL] ~ [LABIAL] pair, [g]_d/ and [d]_g/. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the asymmetrical [DORSAL] ~ [LABIAL] contrast for plosives at the midline electrodes (Fz, Cz, Pz).

5.1.5 Experiment 3: Place & Manner Features: Standards, Deviants, & MMN Waveforms

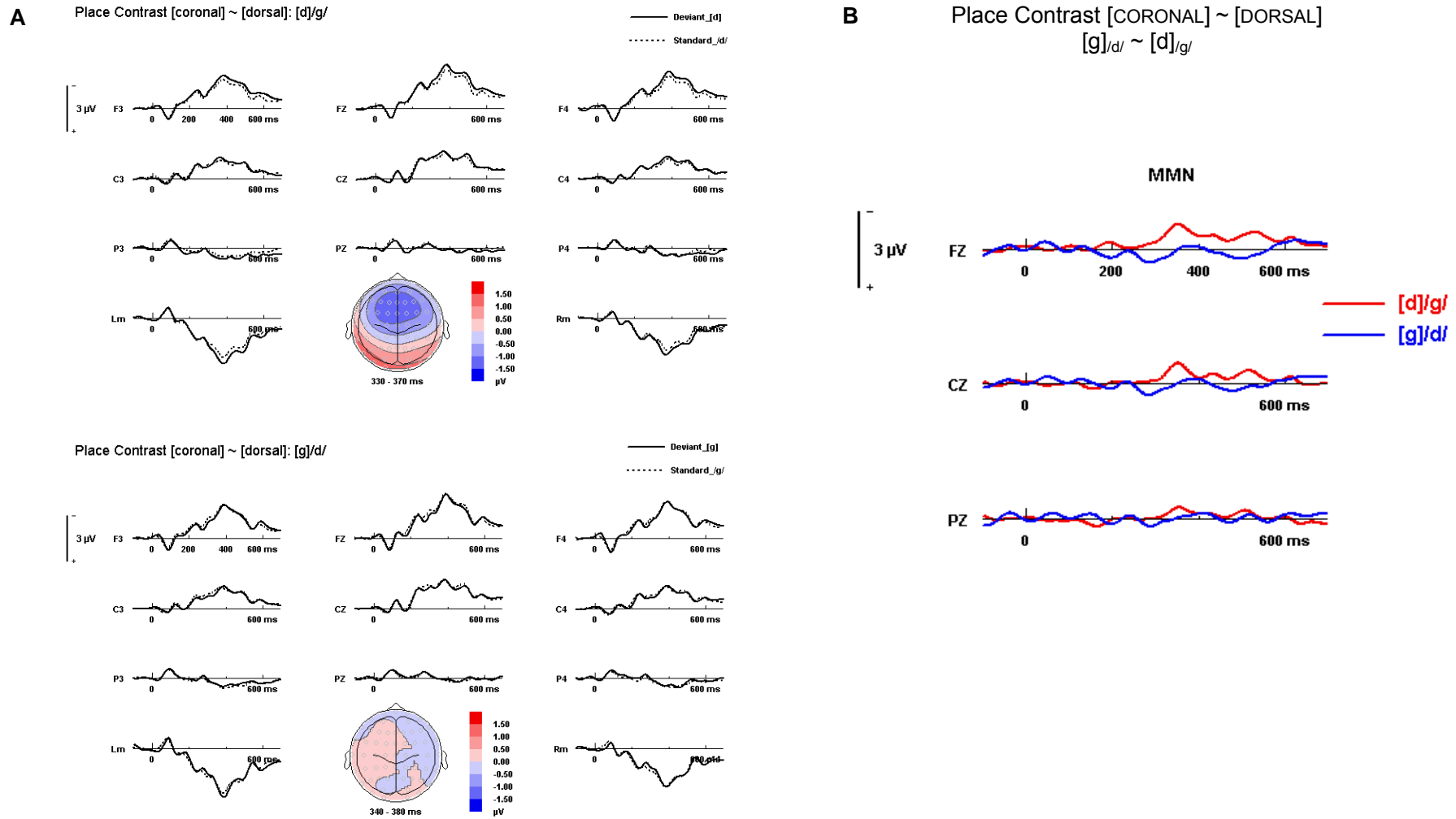


Figure 5.1.5a: Grand-average waveforms for the place contrast [g]_d/ and [d]_g/. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the asymmetrical place contrasts at the midline electrodes (Fz, Cz, Pz).

5.1.5 Experiment 3: Place & Manner Features: Standards, Deviants, & MMN Waveforms

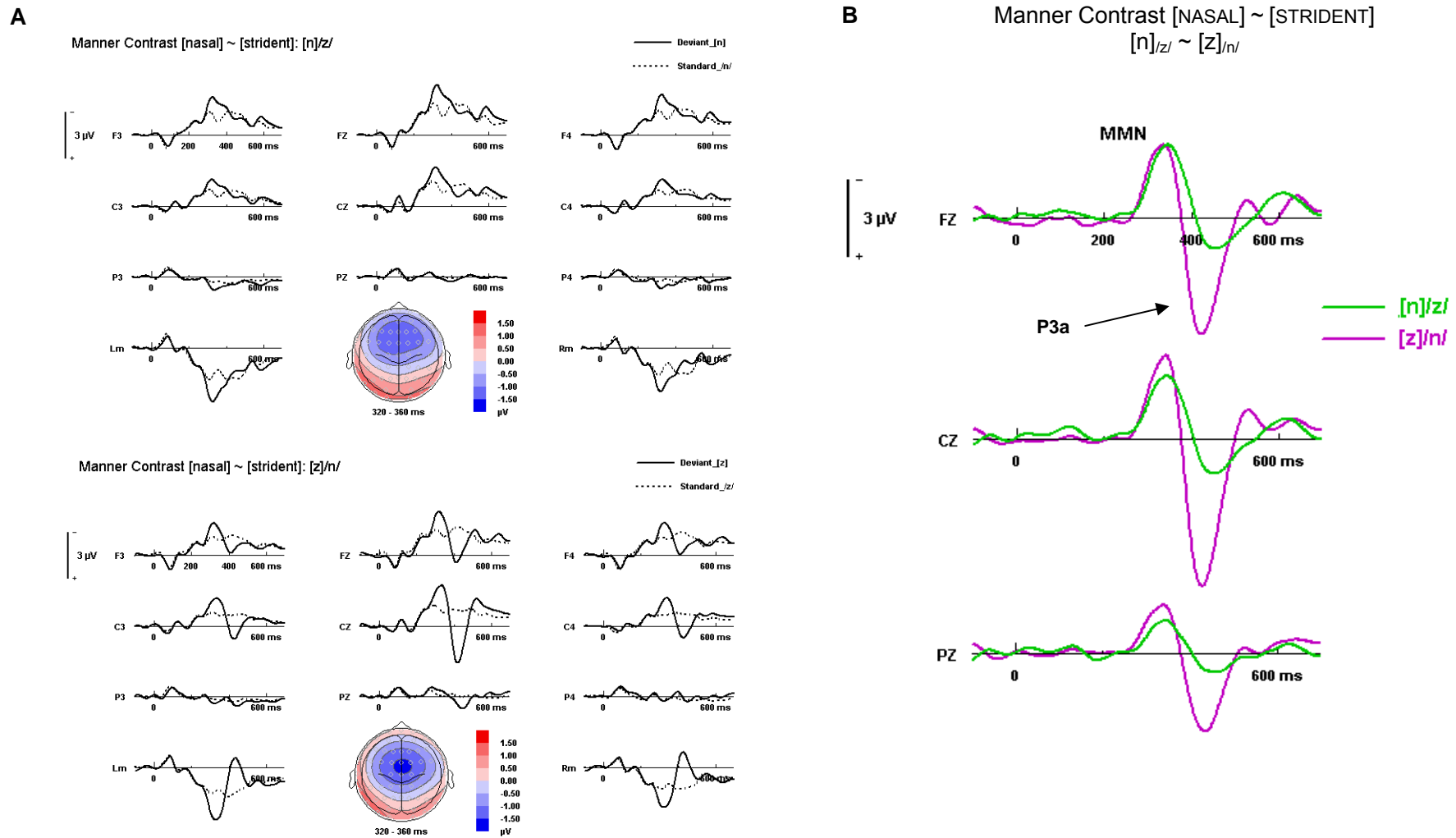


Figure 5.1.5b: Grand-average waveforms for the manner contrast [n]/z/ and [z]/n/. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the symmetrical manner contrast at the midline electrodes (Fz, Cz, Pz). Note the large P3a for the condition [z]/n/.

5.1.6 Experiment 4: Manner Features: Standards, Deviants, & MMN Waveforms

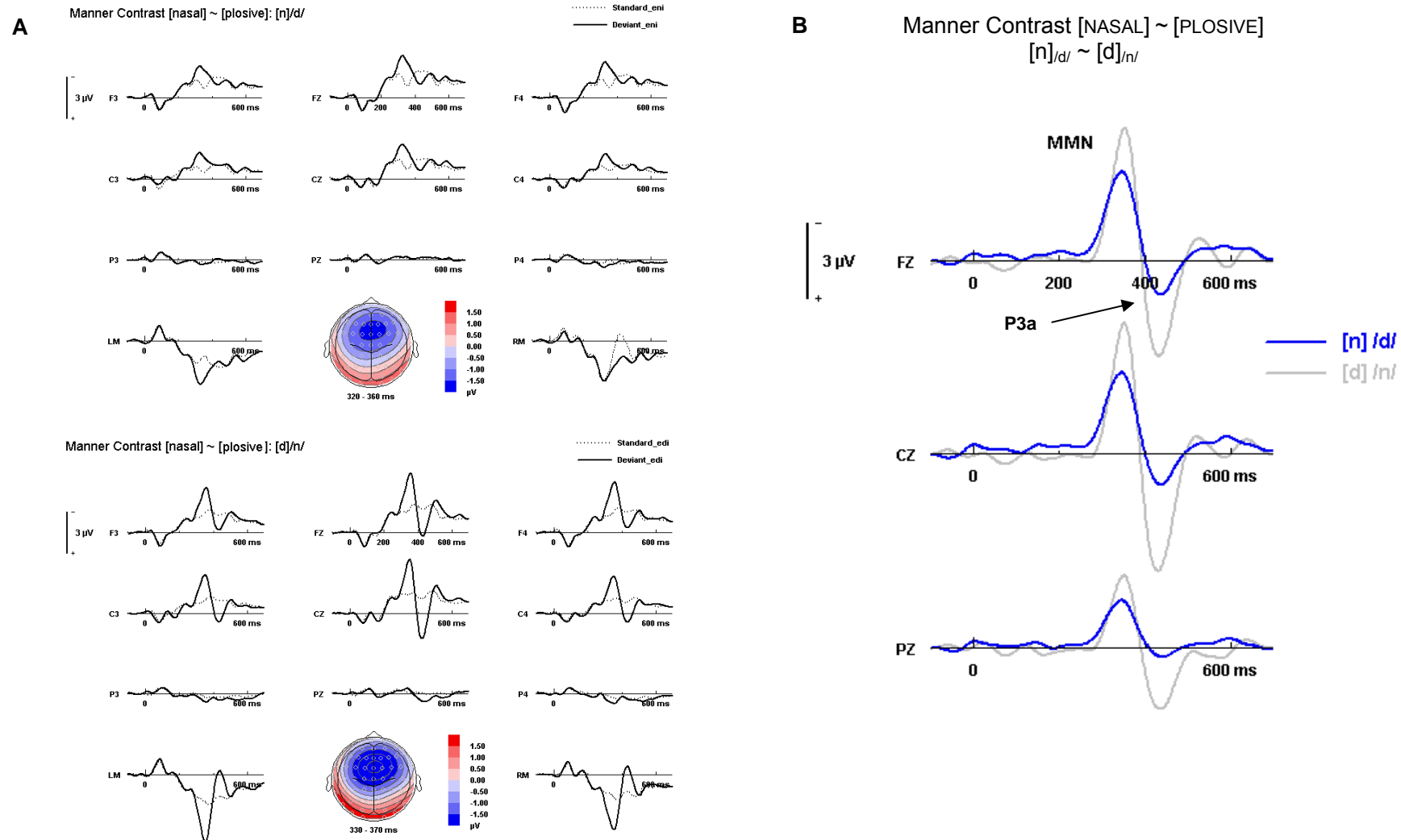


Figure 5.1.6a: Grand-average waveforms for the manner contrast, [n]_{/d/} and [d]_{/n/}. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the asymmetrical manner contrast at the midline electrodes (Fz, Cz, Pz). Note the large P3a for the condition [d]_{/n/}.

5.1.6 Experiment 4: Manner Features: Standards, Deviants, & MMN Waveforms

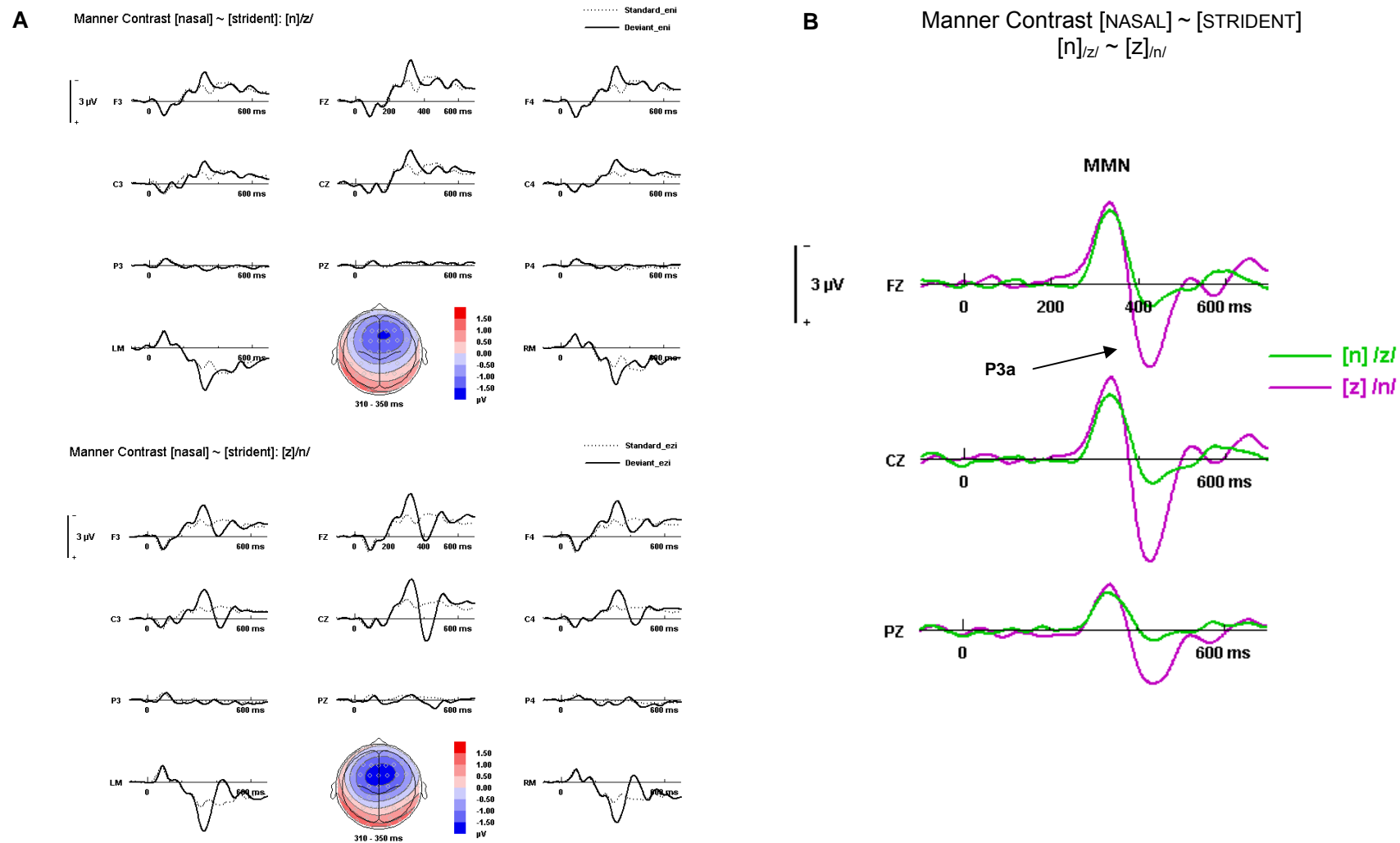


Figure 5.1.6b: Grand-average waveforms for the manner contrast, [n]_{/z/} and [z]_{/n/}. (A) shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). (B) shows the difference wave of the symmetrical manner contrast at the midline electrodes (Fz, Cz, Pz). Note the large P3a for the condition [z]_{/n/}.

5.1.6 Experiment 4: Manner Features: Standards, Deviants, & MMN Waveforms

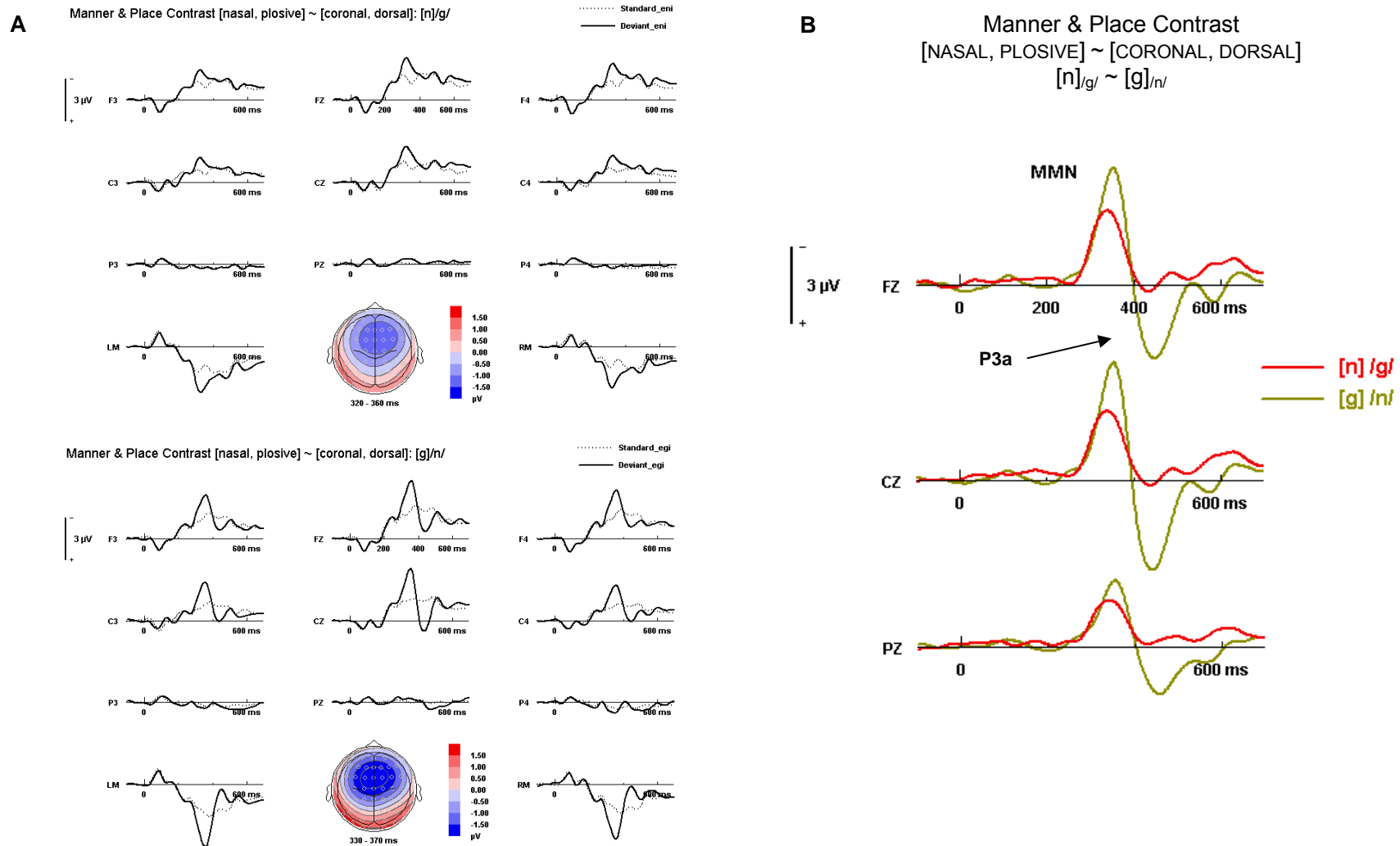


Figure 5.1.6c: Grand-average waveforms for the manner & place contrast, [n]/g/ and [g]/n/. (A) shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). (B) shows the difference wave of the combined manner & place contrast at the midline electrodes (Fz, Cz, Pz). Note the large P3a for the condition [g]/n/.

5.1.7 Experiment 5: Place & Manner Features: Standards, Deviants, & MMN Waveforms

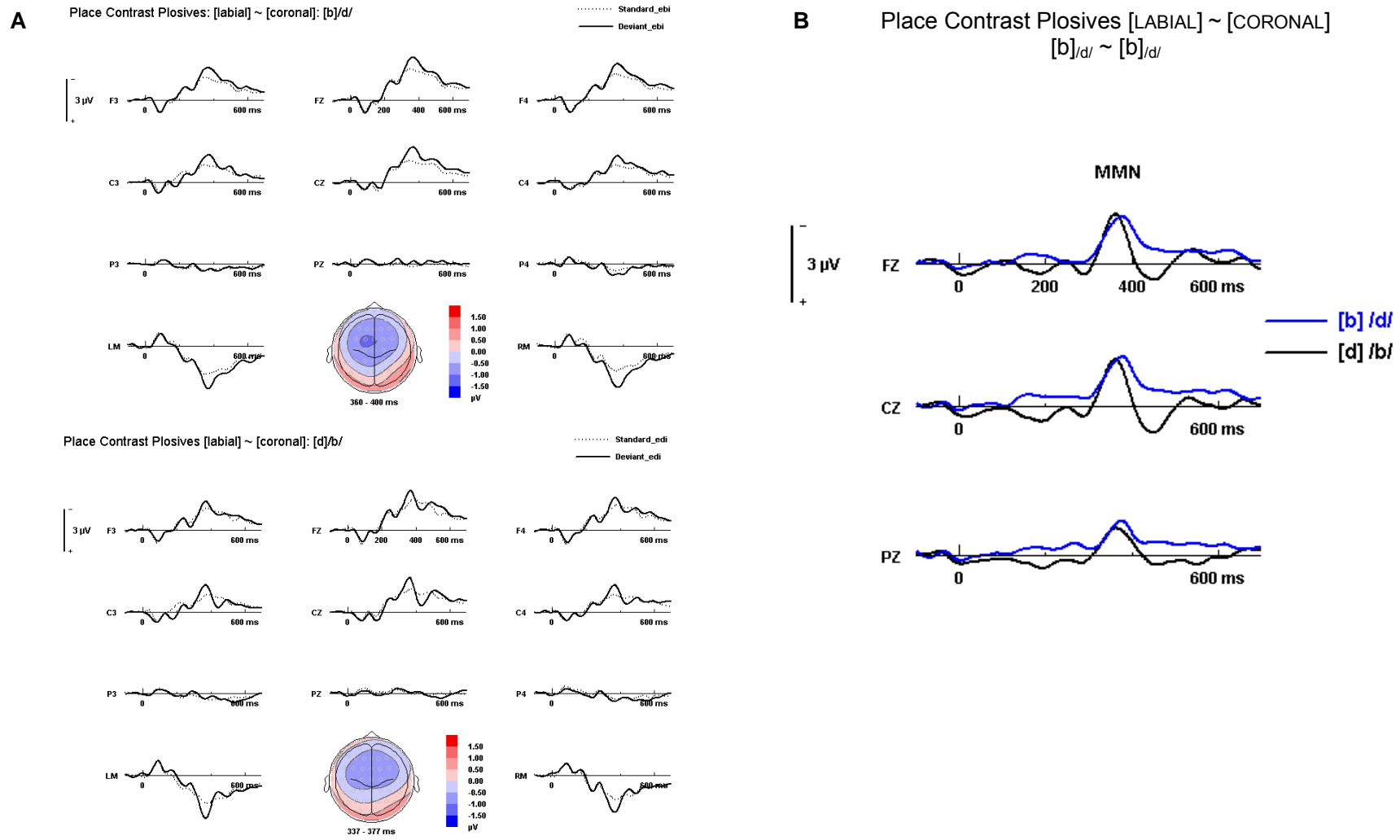


Figure 5.1.7a: Grand-average waveforms for the place contrast with plosives, [b]_{/d/} and [d]_{/b/}. (A) shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). (B) shows the difference wave of the place contrast with plosives at the midline electrodes (Fz, Cz, Pz).

5.1.7 Experiment 5: Place & Manner Features: Standards, Deviants, & MMN Waveforms

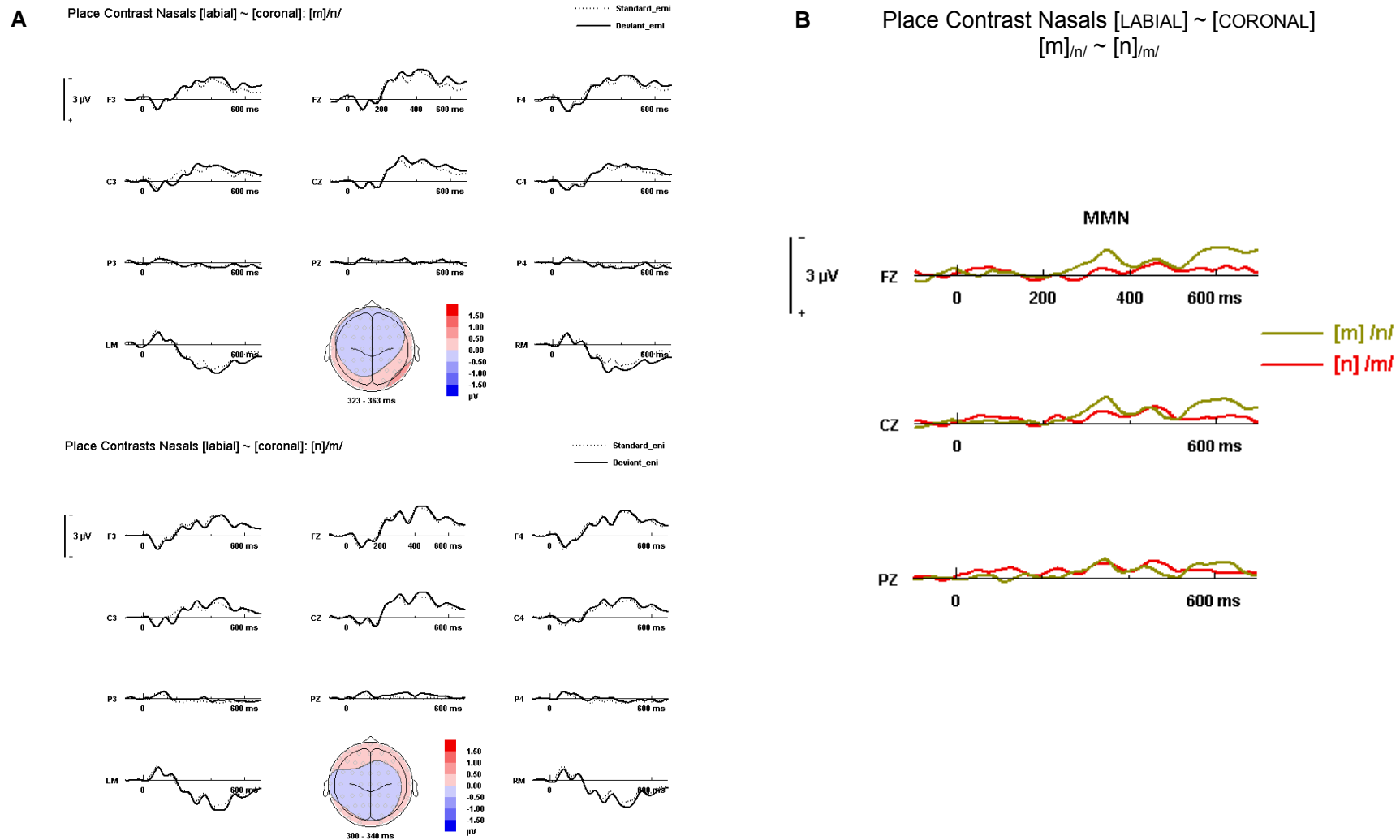


Figure 5.1.7b: Grand-average waveforms for the place contrast with nasals, [m]_{/n/} and [n]_{/m/}. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the place contrast with nasals at the midline electrodes (Fz, Cz, Pz).

5.1.7 Experiment 5: Place & Manner Features: Standards, Deviants, & MMN Waveforms

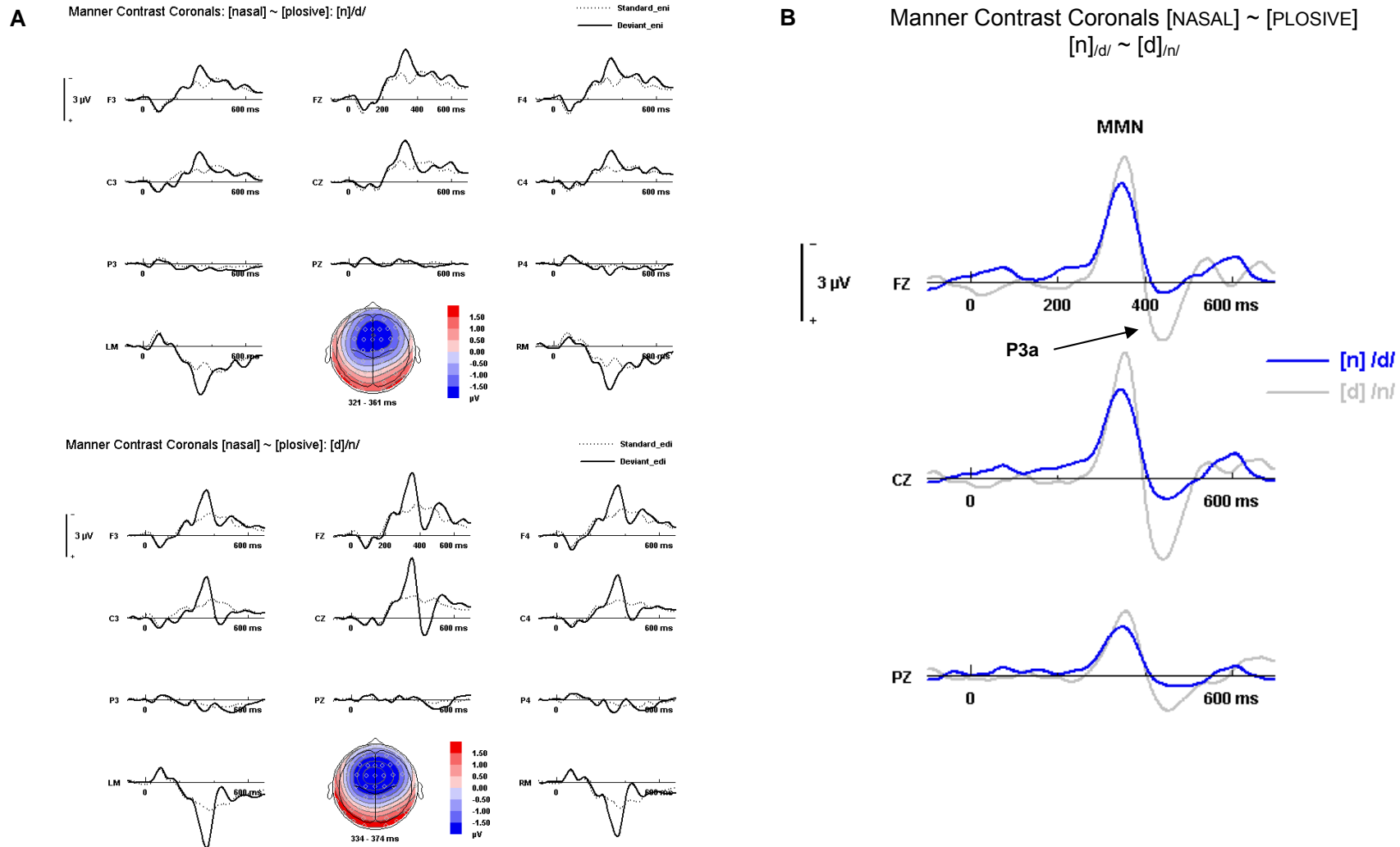


Figure 5.1.7c: Grand-average waveforms for the manner contrast with coronals, [n]_{/d/} and [d]_{/n/}. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the combined manner contrast with coronals at the midline electrodes (Fz, Cz, Pz). Note the large P3a for the condition [d]_{/n/}.

5.1.7 Experiment 5: Place & Manner Features: Standards, Deviants, & MMN Waveforms

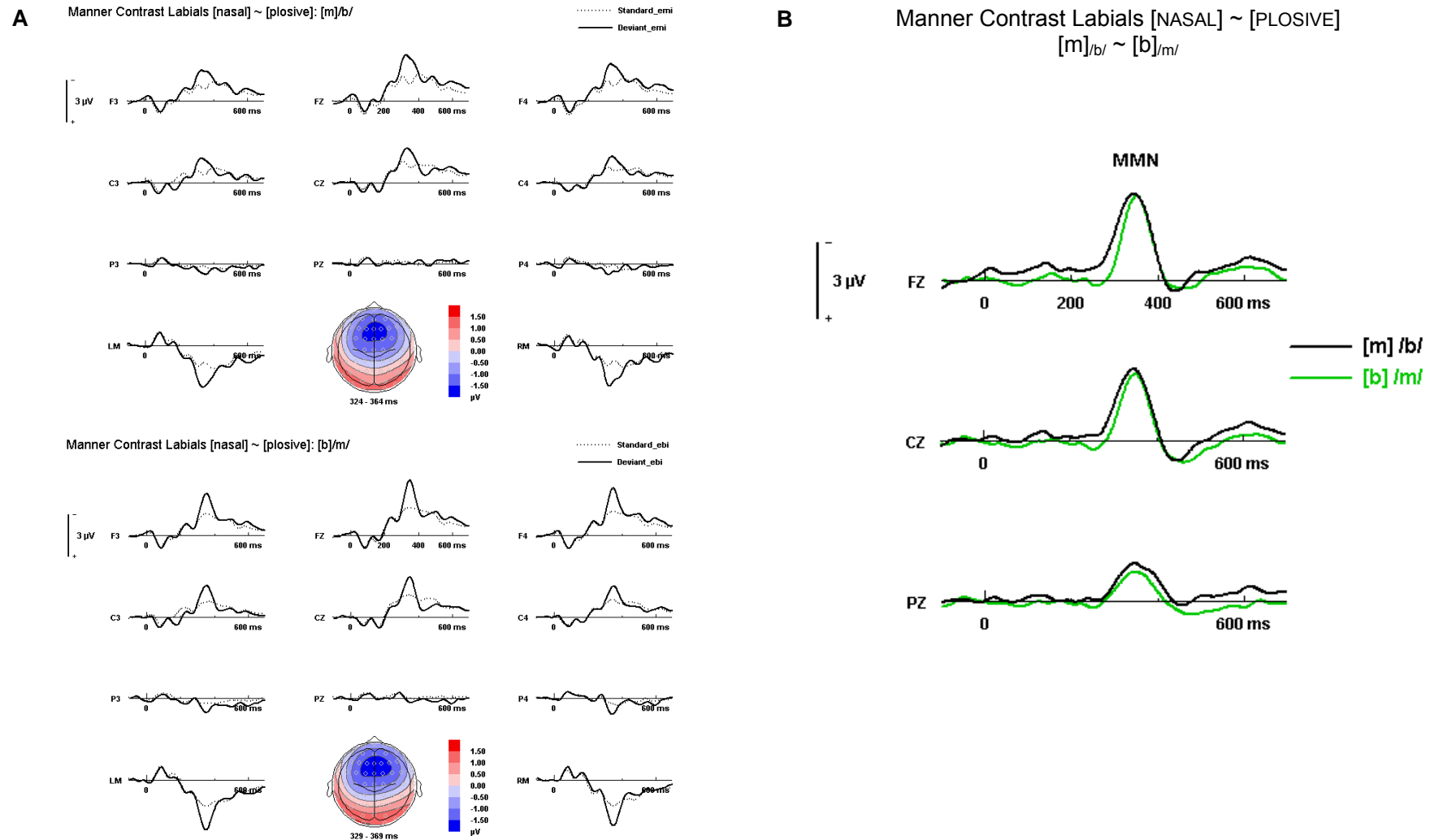


Figure 5.1.7d: Grand-average waveforms for the manner contrast with labials, [m]_{/b/} and [b]_{/m/}. **(A)** shows the standard and deviant waveforms of 11 electrode positions including their topographical maps of the difference waveform (MMN). **(B)** shows the difference wave of the combined manner contrast with labials at the midline electrodes (Fz, Cz, Pz).

5.2 Statistics

5.2.1 Experiment 1: Standard/Deviant Analysis

Not finding a significant interaction with the factor WORDNESS in the MMN time window in either analysis was surprising. To further get an idea about the impact of lexicality onto the automatic processing of our stimuli, we tested for the well known N400 effect which is sensitive to the time course of cognitive processes underlying word comprehension, in particular semantic processing (Kutas & Hillyard, 1980). Although it is especially large to semantic violations, the N400 is not simply an index of violation, but rather a part of the brain's normal response to words that cannot be integrated semantically into the preceding context (in all modalities). The N400 furthermore is related to the processing of isolated words or word-like stimuli, such as pronounceable nonwords (e.g. Picton & Hillyard, 1988; Attias & Praat, 1992; Friedrich et al., 2008). In addition to its sensitivity during language comprehension, the N400 amplitude also appears to be sensitive to the ease of accessing information from long-term memory. It seems to reflect facilitated activation of features of the long-term memory representation that is associated with a lexical item (Kutas & Federmeier, 2000). This view implies that any factor that facilitated lexical access should reduce the N400 amplitude. For instance words with a higher frequency of occurrence result in smaller N400 responses than lower-frequency words, as do repeated words. Similarly, word-like nonwords often elicit larger N400 responses than words (for review see Lau, Phillips & Poeppel, 2008).

This enhanced lexico-semantic memory search for nonwords could have influenced our results beyond our MMN effects. In contrast to words, we expect difficulties in lexical integration of our nonwords, which should be reflected in an enhanced N400 effect.

We restricted this analysis to the mean amplitudes of the standard and deviant ERP data for the Fz²⁶ (to stay consistent with the MMN analysis) in a time window between 350-450 ms, which should well cover the N400 effect. Based on visual inspection this effect was most pronounced at fronto-central electrode sites.

Again a repeated-measures ANOVA for each mean amplitude value with the factors STANDARD-DEVIANT (standard vs. deviant), WORDNESS (word vs. nonword), PAIR OF INVERSION (POI) (non-conflicting pair [e]_{/ø/} ~ [ø]_{/e/} vs. conflicting pair [ø]_{/ø/} ~ [o]_{/ø/}) and DIRECTION OF ACOUSTIC CHANGE (DOC) (F2 falling: [ø]_{/e/} ~ [o]_{/ø/} vs. F2 rising: [e]_{/ø/} ~ [ø]_{/ø/}) was performed.

²⁶ The Fz electrode was used for parameterization to be consistent with the MMN analysis. Usually the N400 is maximal over centro-parietal electrode sites (Cz or Pz). In my data the effects seemed most pronounced at fronto-central sites.

The overall ANOVA revealed a three-way interaction with WORDNESS x POI x DOC ($F(1, 15) = 7.76; p < .014$), a tendency of a two-way interaction STANDAD-DEVIANT x DOC ($F(1, 15) = 4.11; p < .061$) and POI x DOC ($F(1, 15) = 3.52; p < .080$), as well as a main effect for WORDNESS ($F(1, 15) = 13.28; p < .002$), and STANDAD-DEVIANT ($F(1, 15) = 29.17; p < .000$) (for detailed statistical analysis of this standard/deviant analysis see Appendices: 5.2.2 Statistical Table 14d).

Resolving the three-way interaction for WORDNESS, resulted in a significant interaction for POI x DOC ($F(1, 15) = 9.63; p < .007$) for words, but not for the nonwords ($F(1, 15) = 0.06; p > .1; n.s.$). Words showed a significant DOC main effect for the non-conflicting conditions $[e]_{/o/} \sim [\emptyset]_{/e/}$ ($F(1, 15) = 9.15; p < .009$) but not for the conflicting contrast $[\emptyset]_{/o/} \sim [o]_{/o/}$. So far, this analysis revealed a significant difference between words and nonwords, which seems to reflect a N400 effect. This effect mainly is modulated by the words, in particular by the words of the non-conflicting pair of inversion $[e]_{/o/}$ and $[\emptyset]_{/e/}$.

Looking at the standards and deviants separately is licensed by the tendency of the two-way interaction of STANDAD-DEVIANT x DOC. The standards revealed a significant main effect of DOC ($F(1, 15) = 5.42; p < .034$), the deviants however did not ($F(1, 15) = 0.46; p > .1; n.s.$). This means that the difference between the phonological non-conflicting – conflicting conditions are modulated by the standards not by the deviants, which makes sense since the standards are assumed to tap into lexical representations.

The Beginning of a Discussion...

Looking at the amplitude differences beyond the MMN, words clearly differ from nonwords in the N400 time window, which is reflected in the WORDNESS effect. However the effect was reversed in that our words showed a greater negativity than the nonwords at Fz. The N400 for the words seems to be modulated by the conflicting – non-conflicting conditions, in particular by the non-conflicting condition. This was not the case for the nonwords. The same pattern could be seen in the standards (see Figure 5.2.1) as they seem to tap into lexical representations. This was not found for the deviant stimuli. The enhanced amplitudes for our non-conflicting words is surprising as it was expected to find enhanced effects for our nonwords assuming facilitated lexico-semantic memory search and activation of lexical representations. The enhanced effects for the non-conflicting coronal $[e] \sim [\emptyset]$ word conditions could reflect an enhanced lexical search compared to the conflicting $[o] \sim [\emptyset]$ word contrast in that more candidate words might be activated. The enhanced amplitude could mirror facilitated costs since the coronal information remains underspecified. An enhanced search for the best match might be activated in the lexicon more than for the conflicting condition, where there might be a faster and more precise access to stored

information in the mental lexicon. The activation of the specified information might already be channeled and in turn, fewer cohorts are activated, leaving less ambiguity and resulting in reduced ERP activation.

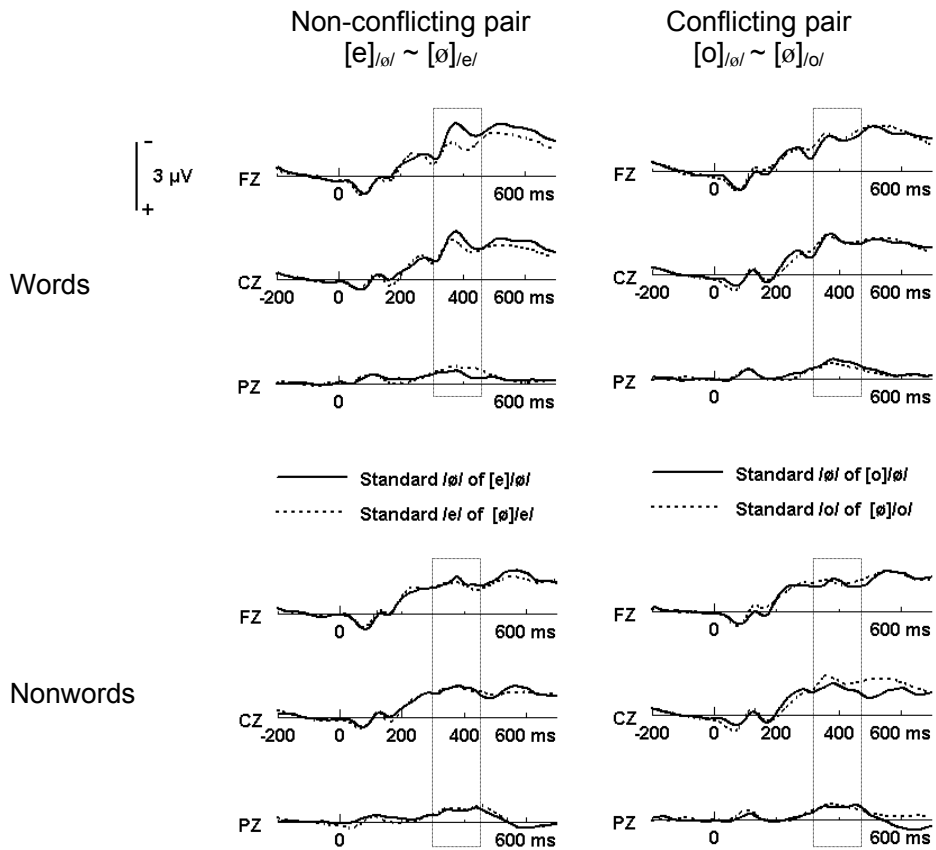


Figure 5.2.1: Grand-average waveforms of the standards for words and nonwords in the conflicting and non-conflicting conditions. The squared areas mark the N400 time window between 350-450 ms. One can clearly see an N400 effect for the words in the non-conflicting vowel contrast.

5.2.2 Statistical Tables

5.2.2.1 Experiment 1: Place Features (Vowels)

Table 14a: Experiment 1: Statistical Analysis for Place Features (Vowels): MMN Mean Amplitudes

MMN Individual Mean Amplitudes (μ V)						
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Wordness</i>						
	words				-2.46	0.38
	nonwords				-2.38	0.33
<i>Pair of Inversion (POI)</i>						
	non-conflict: [e] _{/ø/} - [ø] _{/e/}	15	0.53	.476	-2.47	0.39
	conflict: [o] _{/ø/} - [ø] _{/o/}				-2.36	0.32
<i>Direction of Change</i>						
	F2 falling: [e] _{/ø/} - [ø] _{/o/}	15	3.53	.080	-2.60	0.40
	F2 rising: [ø] _{/e/} - [o] _{/ø/}				-2.23	0.31
<i>Wordness x POI x DOC</i>						
		15	0.27	.611		
<i>POI x DOC</i>						
	non-conflict: [e] _{/ø/} - [ø] _{/e/}	15	6.28	.024*		
	conflict: [o] _{/ø/} - [ø] _{/o/}	15	1.20	.290		
	F2 falling: [e] _{/ø/} - [ø] _{/o/}	15	6.39	.023*		
	F2 rising: [o] _{/ø/} - [ø] _{/e/}	15	3.39	.085		
	F2 rising: [o] _{/ø/} - [ø] _{/e/}	15	5.74	.030*		
<i>Wordness x POI</i>						
		15	0.01	.909		
<i>Wordness x DOC</i>						
		15	0.18	.677		
source	level	df	T-value	p-value	mean	SEM
<i>Words</i>						
non-conflict	F2 falling: MMN_ _{[e]_{/ø/}}	15	-5.58	.000*	-2.39	0.43
	F2 rising: MMN_ _{[ø]_{/e/}}	15	-8.05	.000*	-2.60	0.32
conflict	F2 rising: MMN_ _{[o]_{/ø/}}	15	-6.44	.000*	-1.81	0.28
	F2 falling: MMN_ _{[ø]_{/o/}}	15	-6.30	.000*	-3.02	0.48
	non-conflict: [e] _{/ø/} - [ø] _{/e/}	15	0.47	.646		
	conflict: [o] _{/ø/} - [ø] _{/o/}	15	-1.96	.069 [#]		
<i>Nonwords</i>						
non-conflict	F2 falling: MMN_ _{[e]_{/ø/}}	15	-5.69	.000*	-2.31	0.41
	F2 rising: MMN_ _{[ø]_{/e/}}	15	-6.69	.000*	-2.58	0.39
conflict	F2 rising: MMN_ _{[o]_{/ø/}}	15	-7.87	.000*	-1.92	0.24
	F2 falling: MMN_ _{[ø]_{/o/}}	15	-9.19	.000*	-2.70	0.29
	non-conflict: [e] _{/ø/} - [ø] _{/e/}	15	0.70	.493		
	conflict: [o] _{/ø/} - [ø] _{/o/}	15	-2.06	.057 [#]		

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; [#] $p < .075$

Table 14b: Experiment 1: Statistical Analysis for Place Features (Vowels):
MMN Peak Amplitudes

MMN Individual Peak Amplitudes (μV)						
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Wordness</i>						
	words	15	0.03	.877	-2.66	0.25
	nonwords				-2.62	0.22
<i>Pair of Inversion (POI)</i>						
	non-conflict: [e] _{/ø/} - [ø] _{/e/}	15	0.00	.948	-2.65	0.25
	conflict: [o] _{/ø/} - [ø] _{/o/}				-2.63	0.21
<i>Direction of Change (DOC)</i>						
	F2 falling: [e] _{/ø/} - [ø] _{/e/}	15	1.99	.178	-2.80	0.29
	F2 rising: [ø] _{/e/} - [o] _{/ø/}				-2.48	0.17
<i>Wordness x POI x DOC</i>						
		15	0.24	.631		
<i>POI x DOC</i>						
	non-conflict: [e] _{/ø/} - [ø] _{/e/} DOC	15	5.21	.037*		
	conflict: [o] _{/ø/} - [ø] _{/o/} DOC	15	1.24	.283		
	F2 falling: [e] _{/ø/} - [ø] _{/e/} POI	15	4.87	.043*		
	F2 rising: [o] _{/ø/} - [ø] _{/e/} POI	15	3.28	.090		
		15	3.70	.073 [#]		
<i>Wordness x POI</i>						
		15	0.18	.676		
<i>Wordness x DOC</i>						
		15	0.99	.333		
source	level	df	T-value	p-value	mean	SEM
<i>words</i>						
non-conflict	F2 falling: MMN_ _{[e]_{/ø/}}	15	-5.31	.000*	-2.24	0.42
	F2 rising: MMN_ _{[ø]_{/e/}}	15	-7.96	.000*	-3.00	0.38
conflict	F2 rising: MMN_ _{[o]_{/ø/}}	15	-7.25	.000*	-2.37	0.33
	F2 falling: MMN_ _{[ø]_{/o/}}	15	-6.19	.000*	-3.02	0.49
non-conflict: [e] _{/ø/} - [ø] _{/e/}	DOC	15	1.36	.194		
conflict: [o] _{/ø/} - [ø] _{/o/}	DOC	15	-1.00	.332		
<i>nonwords</i>						
non-conflict	F2 falling: MMN_ _{[e]_{/ø/}}	15	-7.17	.000*	-2.76	0.39
	F2 rising: MMN_ _{[ø]_{/e/}}	15	-6.38	.000*	-2.58	0.40
conflict	F2 rising: MMN_ _{[o]_{/ø/}}	15	-6.19	.000*	-1.97	0.32
	F2 falling: MMN_ _{[ø]_{/o/}}	15	-9.36	.000*	-3.18	0.34
non-conflict: [e] _{/ø/} - [ø] _{/e/}	DOC	15	-0.38	.712		
conflict: [o] _{/ø/} - [ø] _{/o/}	DOC	15	-2.44	.027*		

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; [#] $p < .075$

Table 14c: Experiment 1: Statistical Analysis for Place Features (Vowels):
MMN Peak Latencies

Individual MMN Peak Latencies (ms)						
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Wordness</i>		15	6.14	.026*		
	words				235	2.00
	nonwords				240	2.50
<i>Pair of Inversion (POI)</i>		15	0.32	.577		
	non-conflict: [e] _{/ø/} - [ø] _{/e/}				238	2.68
	conflict: [o] _{/ø/} - [ø] _{/o/}				237	2.16
<i>Direction of Change</i>		15	0.74	.403		
	F2 falling: [e] _{/ø/} - [ø] _{/o/}				237	2.37
	F2 rising: [ø] _{/e/} - [o] _{/ø/}				239	2.42
<i>Wordness x POI x DOC</i>		15	0.15	.703		
<i>POI x DOC</i>		15	2.54	.132		
non-conflict: [e] _{/ø/} - [ø] _{/e/}	DOC	15	1.07	.076		
conflict: [o] _{/ø/} - [ø] _{/o/}	DOC	15	0.41	.530		
F2 falling: [e] _{/ø/} - [ø] _{/o/}	POI	15	0.99	.336		
F2 rising: [o] _{/ø/} - [ø] _{/e/}	POI	15	1.87	.191		
<i>Wordness x POI</i>		15	0.10	.752		
<i>Wordness x DOC</i>		15	0.40	.536		
<i>Words</i>						
non-conflict	F2 falling: MMN_ _{[e]_{/ø/}}				235	2.92
	F2 rising: MMN_ _{[ø]_{/e/}}				238	4.14
conflict	F2 rising: MMN_ _{[o]_{/ø/}}				233	3.28
	F2 falling: MMN_ _{[ø]_{/o/}}				236	2.87
non-conflict: [e] _{/ø/} - [ø] _{/e/}	DOC	15	0.88	.363		
conflict: [o] _{/ø/} - [ø] _{/o/}	DOC	15	0.68	.422		
<i>Nonwords</i>						
non-conflict	F2 falling: MMN_ _{[e]_{/ø/}}				235	4.16
	F2 rising: MMN_ _{[ø]_{/e/}}				245	4.83
conflict	F2 rising: MMN_ _{[o]_{/ø/}}				238	4.72
	F2 falling: MMN_ _{[ø]_{/o/}}				241	4.13
non-conflict: [e] _{/ø/} - [ø] _{/e/}	DOC	15	2.55	.131		
conflict: [o] _{/ø/} - [ø] _{/o/}	DOC	15	0.08	.776		

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

Table 14d: Experiment 1: Statistical Analysis for Place Features (Vowels):
Standard & Deviant Analysis for the N400 Time Window

Standard/ Deviant Analysis [Time Window 350 – 450 ms]						
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Standard-Deviant (StdDev)</i>		15	29.17	.000*		
	standards				-3.13	0.52
	deviants				-4.41	0.50
<i>Wordness</i>		15	13.28	.002*		
	words				-3.37	0.50
	nonwords				-4.16	0.52
<i>Pair of Inversion (POI)</i>		15	0.69	.418		
	non-conflict: [e] _{/ø/} - [ø] _{/e/}				-3.83	0.53
	conflict: [o] _{/ø/} - [ø] _{/o/}				-3.74	0.48
<i>Direction of Change (DOC)</i>		15	0.13	.727		
	F2 falling: [e] _{/ø/} - [ø] _{/o/}				-3.80	0.52
	F2 rising: [ø] _{/e/} - [o] _{/ø/}				-3.77	0.49
<i>StdDev x Wordness</i>		15	0.85	.370		
<i>StdDev x POI</i>		15	1.43	.250		
<i>StdDev x DOC</i>		15	4.11	.061 [#]		
	standards DOC	15	5.42	.034*		
	deviants DOC	15	0.46	.508		
	F2 falling: [e] _{/ø/} - [ø] _{/o/} StdDev	15	12.68	.003*		
	F2 rising: [ø] _{/e/} - [o] _{/ø/} StdDev	15	50.05	.000*		
<i>POI x DOC</i>		15	3.52	.080		
<i>Wordness x POI</i>		15	1.63	.222		
<i>Wordness x DOC</i>		15	0.03	.865		
<i>StdDev x POI x DOC</i>		15	1.45	.248		
<i>StdDev x Wordness x POI</i>		15	0.02	.892		
<i>StdDev x Wordness x DOC</i>		15	0.68	.422		
<i>Wordness x POI x DOC</i>		15	7.76	.014*		
words	POI x DOC	15	9.63	.007*		
	non-conflict: [e] _{/ø/} - [ø] _{/e/} DOC	15	9.15	.009*		
	conflict: [o] _{/ø/} - [ø] _{/o/} DOC	15	2.19	.159		
nonwords	POI x DOC	15	0.06	.808		

* $p < .05$; SEM = Standard Error of Mean; df = degrees of freedom; [#] $p < .075$

5.2.2.2 Experiment 2: Places Features with Plosives

Table 15a: Experiment 2: Statistical Analysis for Place Features (Plosives): Individual MMN Mean Amplitudes of the Early & Late MMN at Fz

MMN Individual Mean Amplitudes (μV)						
Fz electrode						
Early MMN time window [150 - 250 ms]						
source	level	df	F-value	p-value	mean	SEM
Pair of Inversion 1 (POI 1)		13	2.52	.137		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	0.39	.541	-0.62	0.12
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.84	.378	-0.29	0.14
Pair of Inversion 2 (POI 2)		13	0.83	.379		
Asymmetric Dorsal-Coronal pair	[g] _{/d/} - [d] _{/g/}	13	0.84	.377	-0.48	0.13
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.84	.378	-0.29	0.14
Direction of Change 1 (DOC 1)		13	0.04	.852		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	2.42	.144	-0.47	0.10
Back to Front	[b] _{/d/} - [b] _{/g/}	13	0.89	.362	-0.44	0.12
Direction of Change 2 (DOC 2)		13	2.38	.147		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	2.87	.144	-0.41	0.12
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.00	.949	-0.36	0.11
POI 1 x DOC 1		13	0.94	.350		
POI 2 x DOC 2		13	0.11	.741		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13	-3.62	.003*	-0.71	0.20
	MMN [b] _{/d/}	13	-3.03	.010*	-0.53	0.18
Labial-Dorsal	MMN [g] _{/b/}	13	-1.35	.201	-0.23	-0.17
	MMN [b] _{/g/}	13	-2.71	.018*	-0.35	0.13
Dorsal- Coronal	MMN [g] _{/d/}	13	-3.96	.002*	0.59	0.15
	MMN [d] _{/g/}	13	-1.84	.088	0.37	0.20
Late MMN time window [250 - 350 ms]						
source	level	df	F-value	p-value	mean	SEM
Pair of Inversion 1 (POI 1)		13	2.73	.123		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	1.89	.193	-0.59	0.16
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	1.08	.317	-0.29	0.13
Pair of Inversion 2 (POI 2)		13	3.02	.106		
Asymmetric Dorsal-Coronal pair	[g] _{/d/} - [d] _{/g/}	13	1.93	.188	-0.57	0.10
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	1.08	.317	-0.29	0.13
Direction of Change 1 (DOC 1)		13	1.95	.186		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	3.03	.105	-0.60	0.19
Back to Front	[b] _{/d/} - [b] _{/g/}	13	0.57	.462	-0.27	0.13
Direction of Change 2 (DOC 2)		13	2.77	.120		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	4.39	.056*	-0.56	0.14
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.65	.436	-0.30	0.09
POI 1 x DOC 1		13	1.16	.301		
POI 2 x DOC 2		13	0.34	.571		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13	-3.03	.010*	-0.83	0.28
	MMN [b] _{/d/}	13	-1.79	.097	-0.35	0.20
Labial-Dorsal	MMN [g] _{/b/}	13	-2.01	.066 [#]	-0.37	0.19
	MMN [b] _{/g/}	13	-1.62	.129	-0.20	0.12
Dorsal- Coronal	MMN [g] _{/d/}	13	-5.54	.000*	-0.74	0.13
	MMN [d] _{/g/}	13	-2.18	.049*	0.40	0.18

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; [#] $p < .075$

Table 15b: Experiment 2: Statistical Analysis for Place Features (Plosives):
Condition-Related MMN Mean Amplitudes of the Early MMN at Fz and Cz

MMN Condition-Related-Mean Amplitudes (µV) - early MMN						
[b]/d/:	198 – 238 ms	[d]/b/:	178 – 218 ms			
[b]/g/:	173 – 213 ms	[g]/b/:	173 – 213 ms			
[d]/g/:	187 – 227 ms	[g]/d/:	164 – 204 ms			
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Pair of Inversion 1 (POI 1)</i>		13	1.58	.295		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	0.65	.434	-0.29	0.15
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.05	.828	-0.08	0.17
<i>Pair of Inversion 2 (POI 2)</i>		13	0.05	.835		
Asymmetric Dorsal-Coronal pair	[g] _{/d/} - [d] _{/g/}	13	1.38	.261	-0.13	0.17
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.05	.828	-0.08	0.17
<i>Direction of Change 1 (DOC 1)</i>		13	0.31	.495		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	1.00	.335	-0.25	0.14
Back to Front	[b] _{/d/} - [b] _{/g/}	13	0.05	.826	-0.12	0.17
<i>Direction of Change 2 (DOC 2)</i>		13	0.72	.411		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	0.83	.379	-0.20	0.18
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.30	.591	-0.01	0.14
<i>POI 1 x DOC 1</i>		13	0.54	.475		
<i>POI 2 x DOC 2</i>		13	1.58	.231		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13	-1.52	.151	-0.44	0.28
	MMN [b] _{/d/}	13	-0.89	.391	-0.14	0.18
Labial-Dorsal	MMN [g] _{/b/}	13	-0.24	.814	-0.05	0.19
	MMN [b] _{/g/}	13	-0.47	.649	-0.10	0.21
Dorsal- Coronal	MMN [g] _{/d/}	13	-1.25	.234	-0.34	0.28
	MMN [d] _{/g/}	13	0.29	.776	0.08	0.21
Cz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Pair of Inversion 1 (POI 1)</i>		13	6.86	.021*		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	1.64	.223	-0.42	0.18
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.92	.355	0.05	0.13
<i>Pair of Inversion 2 (POI 2)</i>		13	2.02	.179		
Asymmetric Dorsal-Coronal pair	[g] _{/d/} - [d] _{/g/}	13	0.09	.764	-0.18	0.16
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.92	.355	0.05	0.13
<i>Direction of Change 1 (DOC 1)</i>		13	0.35	.564		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	5.95	.030	-0.25	0.16
Back to Front	[b] _{/d/} - [b] _{/g/}	13	0.67	.427	-0.12	0.18
<i>Direction of Change 2 (DOC 2)</i>		13	0.03	.862		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	2.41	.114	-0.05	0.18
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.09	.763	-0.09	0.14
<i>POI 1 x DOC 1</i>		13	2.90	.112		
<i>POI 2 x DOC 2</i>		13	0.64	.437		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13	-2.19	.047*	-0.64	0.29
	MMN [b] _{/d/}	13	-0.90	.385	-0.19	0.21
Labial-Dorsal	MMN [g] _{/b/}	13	1.06	.311	0.15	0.14
	MMN [b] _{/g/}	13	-0.27	.795	-0.05	0.18
Dorsal- Coronal	MMN [g] _{/d/}	13	-0.85	.410	-0.24	0.28
	MMN [d] _{/g/}	13	-0.66	.519	-0.18	0.19

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

Table 15c: Experiment 2: Statistical Analysis for Place Features (Plosives):
Condition-Related MMN Mean Amplitudes of the Late MMN at Fz & Cz

MMN Condition-Related-Mean Amplitudes (μV) - late MMN						
Fz electrode						
[b]/d/:	322 – 362 ms	[d]/b/:	288 – 328 ms			
[b]/g/:	351 – 391 ms	[g]/b/:	253 – 293 ms			
[d]/g/:	269 – 309 ms	[g]/d/:	303 – 343 ms			
source	level	df	F-value	p-value	mean	SEM
<i>Pair of Inversion 1 (POI 1)</i>		13	0.09	.765		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	2.42	.144	-0.25	0.19
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.48	.501	-0.18	0.16
<i>Pair of Inversion 2 (POI 2)</i>		13	0.13	.723		
Asymmetric Dorsal-Coronal	[g] _{/d/} - [d] _{/g/}	13	1.90	.191	-0.26	0.16
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.48	.501	-0.18	0.16
<i>Direction of Change 1 (DOC)</i>		13	0.67	.429		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	1.89	.193	-0.30	0.20
Back to Front	[b] _{/d/} - [b] _{/g/}	13	0.72	.412	-0.13	0.14
<i>Direction of Change 2 (DOC)</i>		13	0.34	.568		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	1.72	.214	-0.27	0.17
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.30	.595	-0.17	0.12
<i>POI 1 x DOC 1</i>		13	4.98	.044*		
<i>POI 2 x DOC 2</i>		13	3.33	.091		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13	-1.88	.083	-0.48	0.25
	MMN [b] _{/d/}	13	-0.11	.915	-0.03	0.26
Labial-Dorsal	MMN [g] _{/b/}	13	-0.57	.579	-0.12	0.21
	MMN [b] _{/g/}	13	-1.71	.110	-0.24	0.14
Dorsal- Coronal	MMN [g] _{/d/}	13	-2.17	.049*	-0.42	0.19
	MMN [d] _{/g/}	13	0.29	.776	-0.10	0.21
Cz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Pair of Inversion 1 (POI 1)</i>		13	0.06	.804		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	3.79	.074 [#]	-0.17	0.14
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.11	.751	-0.20	0.14
<i>Pair of Inversion 2 (POI 2)</i>		13	1.24	.286		
Asymmetric Dorsal-Coronal	[g] _{/d/} - [d] _{/g/}	13	0.06	.811	-0.40	0.17
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.11	.751	-0.20	0.14
<i>Direction of Change 1 (DOC)</i>		13	2.34	.150		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	0.80	.386	-0.35	0.20
Back to Front	[b] _{/d/} - [b] _{/g/}	13	1.77	.206	-0.02	0.14
<i>Direction of Change 2 (DOC)</i>		13	0.58	.461		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	0.38	.547	-0.33	0.16
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.55	.471	-0.26	0.12
<i>POI 1 x DOC 1</i>		13	2.21	.161		
<i>POI 2 x DOC 2</i>		13	0.00	.948		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13	-2.01	.066 [#]	-0.46	0.23
	MMN [b] _{/d/}	13	-0.62	.547	0.13	0.21
Labial-Dorsal	MMN [g] _{/b/}	13	-1.07	.304	-0.25	0.23
	MMN [b] _{/g/}	13	-1.21	.249	-0.16	0.14
Dorsal- Coronal	MMN [g] _{/d/}	13	-2.18	.049*	-0.42	0.19
	MMN [d] _{/g/}	13	-1.70	.114	-0.36	0.21

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; [#] $p < .075$

Table 15d: Experiment 2: Statistical Analysis for Place Features (Plosives) at Fz:
Individual MMN Peak Latencies of the Early and Late MMN

Individual Peak Latencies (ms): Early MMN [150 - 250 ms]						
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
<i>Pair of Inversion 1 (POI 1)</i>		13	0.20	.663		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	0.01	.922	197	11.83
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.39	.545	197	6.28
<i>Pair of Inversion 2 (POI 2)</i>		13	0.01	.908		
Asymmetric Dorsal-Coronal pair	[g] _{/d/} - [d] _{/g/}	13	0.56	.467	201	9.40
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.39	.545	197	6.28
<i>Direction of Change 1 (DOC 1)</i>		13	0.14	.719		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	0.02	.900	197	6.16
Back to Front	[b] _{/d/} - [b] _{/g/}	13	0.32	.581	200	5.69
<i>Direction of Change 2 (DOC 2)</i>		13	0.04	.843		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	0.75	.402	201	6.10
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.12	.733	200	5.76
<i>POI 1 x DOC 1</i>		13	0.10	.757		
<i>POI 2 x DOC 2</i>		13	0.65	.436		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13			196	10.31
	MMN [b] _{/d/}	13			197	8.11
Labial-Dorsal	MMN [g] _{/b/}	13			197	5.12
	MMN [b] _{/g/}	13			203	6.88
Dorsal- Coronal	MMN [g] _{/d/}	13			205	9.38
	MMN [d] _{/g/}	13			197	11.83
Individual Peak Latencies (ms): Late MMN [250 - 350 ms]						
Fz electrode						
<i>Pair of Inversion 1 (POI 1)</i>		13	1.97	.184		
Asymmetric Labial-Coronal pair	[d] _{/b/} - [b] _{/d/}	13	0.31	.587	303	5.95
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.48	.501	294	5.88
<i>Pair of Inversion 2 (POI 2)</i>		13	0.00	.973		
Asymmetric Dorsal-Coronal pair	[g] _{/d/} - [d] _{/g/}	13	1.08	.318	294	5.61
Symmetric Labial-Dorsal pair	[g] _{/b/} - [b] _{/g/}	13	0.48	.501	294	5.88
<i>Direction of Change 1 (DOC 1)</i>		13	0.02	.896		
Front to Back	[d] _{/b/} - [g] _{/b/}	13	3.05	.104	298	5.61
Back to Front	[b] _{/d/} - [b] _{/g/}	13	0.04	.843	299	7.11
<i>Direction of Change 2 (DOC 2)</i>			0.06	.813		
Front to Back	[g] _{/b/} - [g] _{/d/}	13	1.21	.292	295	6.25
Back to Front	[b] _{/g/} - [d] _{/g/}	13	0.79	.390	293	5.75
<i>POI 1 x DOC 1</i>		13	0.87	.368		
<i>POI 2 x DOC 2</i>		13	1.92	.190		
Null-hypothesis:						
Labial-Coronal	MMN [d] _{/b/}	13			306	7.36
	MMN [b] _{/d/}	13			300	8.33
Labial-Dorsal	MMN [g] _{/b/}	13			290	7.00
	MMN [b] _{/g/}	13			298	9.20
Dorsal- Coronal	MMN [g] _{/d/}	13			300	8.29
	MMN [d] _{/g/}	13			288	6.97

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

5.2.2.3 Experiment 3: Places & Manner Features

Table 16a: Experiment 3: Statistical Analysis for Place & Manner Features for Individual MMN Mean Amplitudes at Fz, & Condition-Related MMN Mean Amplitudes at Fz & Cz

MMN Individual Mean Amplitudes (μV)						
Fz electrode						
	level	df	T-value	p-value	mean	SEM
null-hypothesis	MMN [n]/z/	25	-10.38	.000*	-3.03	0.29
	MMN [z]/n/	25	-10.18	.000*	-3.17	0.31
	MMN [d]/g/	25	-5.37	.000*	-1.29	0.24
	MMN [g]/d/	25	-2.13	.043*	-0.49	0.23
Contrasts	manner change [n]/z/ ~ [z]/n/	25	-0.43	.669	-3.10	0.25
	place change [d]/g/ ~ [g]/d/	25	5.55	.017*	-0.89	0.17
	difference manner - place	25	2.35	.026*	-0.47	0.20
MMN Condition-Related-Mean Amplitudes [μV] (40 ms)						
MMN [n]/z/: 322 - 362 ms; MMN [z]/n/: 310 - 350 ms						
MMN [d]/g/: 330 - 360 ms; MMN [g]/d/: 344 - 384 ms						
Fz electrode						
	level	df	T-value	p-value	mean	SEM
Null-hypothesis	MMN [n]/z/	25	-8.56	.000*	-2.63	0.30
	MMN [z]/n/	25	-8.80	.000*	-2.71	0.32
	MMN [d]/g/	25	-3.81	.001*	-0.93	0.25
	MMN [g]/d/	25	-0.24	.810	-0.07	0.27
Contrasts	manner change [n]/z/ ~ [z]/n/	25	0.26	.816	-2.67	0.26
	place change [d]/g/ ~ [g]/d/	25	2.61	.015*	-0.50	0.20
	difference manner - place	25	-2.01	.055 [#]	-0.39	0.20
Cz electrode						
Null-hypothesis	MMN [n]/z/	25	-8.06	.000*	-2.35	0.29
	MMN [z]/n/	25	-10.58	.000*	-2.89	0.27
	MMN [d]/g/	25	-3.66	.001*	-0.75	0.20
	MMN [g]/d/	25	-0.37	.717	-0.09	0.25
Contrasts	manner change [n]/z/ ~ [z]/n/	25	-1.61	.119	-2.62	0.23
	place change [d]/g/ ~ [g]/d/	25	2.20	.037*	-0.42	0.17
	difference manner - place	25	-2.84	.009*	-1.52	0.17

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; [#] $p < .075$

Table 16b: Experiment 3: Statistical Analysis for Place & Manner Features for Individual MMN Peak Latencies at Fz

Individual Peak Latencies (ms)						
Fz electrode						
source	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	MMN [n]/z/	25			349	4.36
	MMN [z]/n/	25			337	4.07
	MMN [d]/g/	25			357	5.27
	MMN [g]/d/	25			367	5.87
<i>Contrasts</i>	manner change	25	-1.93	.065 [#]	343	2.72
	place change	25	1.58	.127	362	4.50
	difference manner	25	-0.20	.841	11	4.27

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; [#] $p < .075$

Table 16c: Experiment 3: Statistical Analysis for Place & Manner Features at Fz for the P3a

P300 Grand-average Time Window (400-450 ms)						
Fz electrode						
source	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	P300 [n]/z/	25	2.21	.037*	0.57	0.26
	P300 [z]/n/	25	8.45	.000*	4.14	0.49
	P300 [d]/g/	25	-2.21	.036* ¹	-0.52	0.24
	P300 [g]/d/	25	0.25	.809	0.06	0.26
<i>Contrasts</i>	manner change	25	6.24	.000*	2.36	0.27
	place change	25	1.61	.121	-0.23	0.17

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; [#] $p < .075$

¹Note, the significance of [d]/g/ is a negativity, no a positivity.

5.2.2.4 Experiment 4: Manner Features

Table 17a: Experiment 4: Statistical Analysis for Manner Features:
Individual MMN Mean Amplitudes at Fz, & Condition-Related Mean Amplitudes at Fz & Cz

MMN Individual Mean Amplitudes (μV)						
Fz electrode						
source	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	MMN [n]/z/	24	-8.79	.000*	-3.14	0.36
	MMN [z]/n/	24	-8.77	.000*	-3.39	0.39
	MMN [n]/d/	24	-9.78	.000*	-3.55	0.36
	MMN [d]/n/	24	-11.24	.000*	-5.03	0.45
	MMN [n]/g/	24	-9.24	.000*	-3.08	0.33
	MMN [g]/n/	24	-11.96	.000*	-4.36	0.36
<i>Contrasts</i>	manner contrast (control) [Nasal] - [Strident]	24	0.54	.598	-3.27	0.29
	manner contrast [Nasal] - [Plosive]	24	3.05	.006*	-4.29	0.33
	manner & place contrast [Nasal, Coronal] - [Plosive, Dorsal]	24	3.30	.003*	-3.72	0.29
	difference [manner & place] - [manner]	24	-0.41	.683	-1.38	0.36
	difference [manner & place] - [manner (control)]	24	1.83	.080	-0.76	0.33
	difference [manner] - [manner (control)]	24	2.28	.032*	-0.87	0.39

MMN Condition-Related-Mean Amplitudes [μV] (40 ms)						
MMN [n]/z/: 312 - 352 ms; MMN [z]/n/: 314 - 354 ms						
MMN [n]/d/: 325 - 365 ms; MMN [d]/n/: 335 - 375 ms						
MMN [n]/g/: 314 - 354 ms; MMN [g]/n/: 335 - 375 ms						
Fz electrode						
	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	MMN [n]/z/	24	-8.06	.000*	-2.89	0.36
	MMN [z]/n/	24	-8.17	.000*	-3.21	0.39
	MMN [n]/d/	24	-9.04	.000*	-3.36	0.37
	MMN [d]/n/	24	-11.00	.000*	-4.89	0.44
	MMN [n]/g/	24	-7.67	.000*	-2.73	0.35
	MMN [g]/n/	24	-11.46	.000*	-4.28	0.37
<i>Contrasts</i>	manner contrast (control) [Nasal] - [Strident]	24	-0.69	.492	-3.05	0.30
	manner contrast [Nasal] - [Plosive]	24	-3.29	.003*	-4.13	0.33
	manner & place contrast [Nasal, Coronal] - [Plosive, Dorsal]	24	-3.85	.001*	-3.51	0.30
	difference [manner & place] - [manner]	24	0.03	.975	-1.62	0.36
	difference [manner & place] - [manner (control)]	24	2.10	.047*	-0.93	0.31
	difference [manner] - [manner (control)]	24	2.21	.037*	-0.92	0.37

Continues on next page...

Cz electrode						
	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	MMN [n]/z/	24	-8.09	.000*	-2.51	0.31
	MMN [z]/n/	24	-9.39	.000*	-3.19	0.34
	MMN [n]/d/	24	-8.39	.000*	-3.06	0.36
	MMN [d]/n/	24	-13.76	.000*	-4.77	0.35
	MMN [n]/g/	24	-9.29	.000*	-2.56	0.28
	MMN [g]/n/	24	-14.28	.000*	-4.33	0.30
<i>Contrasts</i>	manner contrast (control) [Nasal] - [Strident]	24	-1.48	.153	-2.85	0.23
	manner contrast [Nasal] - [Plosive]	24	-4.08	.000*	-3.92	0.29
	manner & place contrast [Nasal, Coronal] - [Plosive, Dorsal]	24	-5.01	.000*	-3.45	0.23
	difference [manner & place] - [manner]	24	0.11	.912	-1.77	0.31
	difference [manner & place] - [manner (control)]	24	1.83	.079	-1.22	0.28
	difference [manner] - [manner (control)]	24	1.75	.093	-1.19	0.33

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

Table 17b: Experiment 4: Statistical Analysis for Manner Features: Individual MMN Peak Latencies

Individual Peak Latencies (ms)							
Fz electrode							
	source	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	MMN [n]/z/		24			338	3.68
	MMN [z]/n/		24			338	3.24
	MMN [n]/d/		24			344	3.87
	MMN [d]/n/		24			352	2.37
	MMN [n]/g/		24			340	4.00
	MMN [g]/n/		24			354	1.94
<i>Contrasts</i>	manner contrast (control) [Nasal] - [Strident]		24	0.13	.901	338	2.78
	manner contrast [Nasal] - [Plosive]		24	-2.62	.015*	348	2.77
	manner & place contrast [Nasal, Coronal] - [Plosive, Dorsal]		24	-3.12	.005*	347	2.19
	difference [manner & place] - [manner]		24	1.03	.315	11	2.72
	difference [manner & place] - [manner (control)]		24	2.81	.010*	7	3.38
	difference [manner] - [manner (control)]		24	1.95	.063#	4	2.83

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

Table 17c: Experiment 4: Statistical Analysis for Manner Features at Fz for the P3a

P300 Grand-average Time Window (400-450 ms)						
Fz electrode						
source	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	P300 [n]/z/	24	2.25	.034*	0.86	0.38
	P300 [z]/n/	24	7.09	.000*	3.21	0.45
	P300 [n]/d/	24	2.46	.022*	0.99	0.41
	P300 [d]/n/	24	7.14	.000*	3.21	0.45
	P300 [n]/g/	24	0.21	.834	0.07	0.34
	P300 [g]/n/	24	4.98	.000*	2.06	0.41
<i>Contrasts</i>	manner contrast (control) [Nasal] - [Strident]	24	-4.53	.000*	2.03	0.33
	manner contrast [Nasal] - [Plosive]	24	-4.25	.000*	2.10	0.34
	manner & place contrast [Nasal, Coronal] - [Plosive, Dorsal]	24	-3.84	.001*	1.07	0.28

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

5.2.2.5 Experiment 5: Place & Manner Features

Table 18a: Experiment 5: Statistical Analysis for Place and Manner Features:
Individual MMN Mean Amplitudes at Fz, & Condition-Related Mean Amplitudes at Fz & Cz

MMN Individual Mean Amplitudes (μV)						
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
Place Contrast						
<i>Feature-type</i>		23	27.43	.000*		
Nasals	[m]/n/ - [n]/m/	23	0.06	.809	-0.83	0.18
Plosives	[d]/b/ - [b]/d/	23	0.51	.482	-2.03	0.27
<i>Level of Conflict</i>		23	0.16	.692		
Non-conflict	[m]/n/ - [b]/d/	23	11.14	.003*	-1.39	0.22
Conflict	[n]/m/ - [d]/b/	23	13.72	.001*	-1.47	0.24
<i>Feature-type x Level</i>		23	0.43	.519		
Manner Contrast						
<i>Feature-type</i>		23	12.36	.002*		
Coronals	[n]/d/ - [d]/n/	23	12.64	.002*	-4.37	0.47
Labials	[b]/m/ - [m]/b/	23	0.01	.912	-3.48	0.33
<i>Level of Conflict</i>		23	5.24	.032*		
Non-conflict	[n]/d/ - [m]/b/	23	0.85	.366	-3.62	0.39
Conflict	[d]/n/ - [b]/m/	23	11.87	.002*	-4.23	0.43
<i>Feature-type x Level</i>		23	6.09	.021*		

Continues on next page...

	source	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>		MMN_[n]/mʹ	23	-4.00	.001*	-0.79	0.20
		MMN_[m]/n/	23	-3.15	.005*	-0.87	0.28
		MMN_[d]/b/	23	-5.74	.000*	-2.15	0.37
		MMN_[b]/d/	23	-7.35	.000*	-1.90	0.26
		MMN_[n]/d/	23	-8.41	.000*	-3.74	0.44
		MMN_[d]/n/	23	-8.98	.000*	-5.00	0.56
		MMN_[m]/b/	23	-9.20	.000*	-3.50	0.38
		MMN_[b]/mʹ	23	-8.81	.000*	-3.46	0.39

MMN Condition-Related-Mean Amplitudes

Fz electrode

Place Contrasts:

MMN: [n]/mʹ: 304-344 ms

MMN: [m]/n/: 326-366 ms

MMN: [d]/b/: 338-378 ms

MMN: [b]/d/: 354-394 ms

Manner Contrasts:

MMN: [n]/d/: 320-360 ms

MMN: [d]/n/: 334-374 ms

MMN: [m]/b/: 320-360 ms

MMN: [b]/mʹ: 332-372 ms

	source	level	df	F-value	p-value	mean	SEM
Place contrast							
<i>Feature-type</i>							
			23	53.97	.000*		
	Nasals	[m]/n/ -[n]/mʹ	23	1.75	.199	-0.40	0.17
	Plosives	[d]/b/ -[b]/d/	23	0.24	.629	-1.77	0.25
<i>Level of Conflict</i>							
			23	0.26	.613		
	Non-conflict	[m]/n/ -[b]/d/	23	9.54	.005*	-1.16	0.23
	Conflict	[n]/mʹ -[d]/b/	23	25.89	.000*	-1.01	0.25
<i>Feature-type x Level of</i>							
			23	1.62	.216		
Manner contrast							
<i>Feature-type</i>							
			23	11.79	.002*		
	Coronals	[n]/d/ -[d]/n/	23	10.93	.003*	-3.97	0.48
	Labials	[b]/mʹ -[m]/b/	23	0.13	.726	-2.98	0.33
<i>Level of Conflict</i>							
			23	4.35	.048*		
	Non-conflict	[n]/d/ -[m]/b/	23	1.80	.193	-3.22	0.40
	Conflict	[d]/n/ -[b]/mʹ	23	11.31	.003*	-3.73	0.40
<i>Feature-type x Level of</i>							
			23	5.93	.023*		
	<i>Null-hypothesis</i>						
		MMN_[n]/mʹ	23	-0.66	.515	-0.15	0.22
		MMN_[m]/n/	23	-2.33	.029*	-0.65	0.28
		MMN_[d]/b/	23	-5.24	.000*	-1.84	0.36
		MMN_[b]/d/	23	-5.71	.000*	-1.67	0.29
		MMN_[n]/d/	23	-7.19	.000*	-3.40	0.47
		MMN_[d]/n/	23	-8.43	.000*	-4.54	0.54
		MMN_[m]/b/	23	-8.25	.000*	-3.05	0.37
		MMN_[b]/mʹ	23	-7.51	.000*	-2.91	0.39

Cz electrode

	source	level	df	F-value	p-value	mean	SEM
Place contrast							
<i>Feature-type</i>							
			23	66.20	.000*		
	Nasals	[m]/n/ -[n]/mʹ	23	1.03	.321	-0.53	0.15
	Plosives	[d]/b/ -[b]/d/	23	0.12	.735	-1.77	0.25

Continues on next page...

source	level	df	F-value	p-value	mean	SEM
<i>Level of Conflict</i>		23	0.17	.689		
Non-conflict	[m] _{/n/} -[b] _{/d/}	23	13.20	.001*	-1.20	0.23
Conflict	[n] _{/m/} -[d] _{/b/}	23	25.88	.000*	-1.10	0.22
<i>Feature-type x Level of Conflict</i>		23	0.83	.371		
Manner contrast						
<i>Feature-type</i>		23	28.07	.000*		
Coronals	[n] _{/d/} -[d] _{/n/}	23	28.99	.000*	-3.82	0.41
Labials	[b] _{/m/} -[m] _{/b/}	23	0.31	.589	-2.44	0.24
<i>Level of Conflict</i>		23	14.44	.001*		
Non-conflict	[n] _{/d/} -[m] _{/b/}	23	3.16	.089	-2.79	0.32
Conflict	[d] _{/n/} -[b] _{/m/}	23	32.32	.000*	-3.48	0.33
<i>Feature-type x Level of Conflict</i>		23	13.63	.001*		
source	level	df	T-value	p-value	mean	SEM
<i>Null-hypothesis</i>	MMN_ _{[n]_{/m/}}	23	-2.10	.047*	-.37	0.18
	MMN_ _{[m]_{/n/}}	23	-2.80	.000*	-.69	0.25
	MMN_ _{[d]_{/b/}}	23	-5.72	.000*	-1.83	0.32
	MMN_ _{[b]_{/d/}}	23	-5.94	.000*	-1.70	0.29
	MMN_ _{[n]_{/d/}}	23	-7.44	.000*	-3.05	0.41
	MMN_ _{[d]_{/n/}}	23	-9.91	.000*	-4.59	0.46
	MMN_ _{[m]_{/b/}}	23	-8.91	.000*	-2.53	0.28
	MMN_ _{[b]_{/m/}}	23	-8.18	.000*	-2.36	0.28

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

Table 18b: Experiment 5: Statistical Analysis for Place & Manner Features: MMN Peak Latencies at Fz

Individual Peak Latencies (ms)						
Fz electrode						
source	level	df	F-value	p-value	mean	SEM
Place Contrast						
<i>Feature-type</i>		23	11.28	.003*		
Nasals	[m] _{/n/} -[n] _{/m/}	23	5.92	.023*	346	7.08
Plosives	[d] _{/b/} -[b] _{/d/}	23	12.08	.002*	372	3.21
<i>Level of Conflict</i>		23	19.58	.000*		
Non-conflict	[m] _{/n/} -[b] _{/d/}	23	8.93	.007*	365	4.40
Conflict	[n] _{/m/} -[d] _{/b/}	23	10.85	.003*	353	3.88
<i>Feature-type x Level of Conflict</i>		23	0.05	.818		
Manner Contrast						
<i>Feature-type</i>		23	2.00	.171		
Coronals	[n] _{/d/} -[d] _{/n/}	23	0.55	.467	348	2.56
Labials	[b] _{/m/} -[m] _{/b/}	23	0.33	.574	352	2.98
<i>Level of Conflict</i>		23	0.98	.334		
Non-conflict	[n] _{/d/} -[m] _{/b/}	23	0.60	.447	349	3.39
Conflict	[d] _{/n/} -[b] _{/m/}	23	0.62	.439	352	2.47
<i>Feature-type x Level of Conflict</i>		23	0.01	.914		

Continues on next page...

	level	df	T-value	P-value	mean	SEM
<i>Null-hypothesis</i>	MMN [n]/m/	23			341	7.10
	MMN [m]/n/	23			352	7.72
	MMN [d]/b/	23			366	2.93
	MMN [b]/d/	23			378	4.27
	MMN [n]/d/	23			347	3.53
	MMN [d]/n/	23			351	3.32
	MMN [m]/b/	23			351	4.58
	MMN [b]/m/	23			353	2.73

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

Table 18c: Experiment 5: Statistical Analysis for Place & Manner Features at Fz for the P3a

P300 Grand-average Time Window (400–450 ms)						
Fz electrode						
source	level	df	T-value	P-value	mean	SEM
<i>Null-hypothesis</i>	P300 [n]/m/	23	-0.99	.331	-0.22	0.23
	P300 [m]/n/	23	-0.54	.593	-0.11	0.19
	P300 [d]/b/	23	1.15	.261	0.35	0.30
	P300 [b]/d/	23	-3.30	.003* ²	-0.65	0.20
	P300 [n]/d/	23	1.19	.247	0.50	0.42
	P300 [d]/n/	23	3.58	.002*	2.20	0.61
	P300 [m]/b/	23	1.39	.178	0.57	0.41
	P300 [b]/m/	23	0.65	.520	0.31	0.47
<i>Place Contrasts</i>	NASALS: [m]/n/ -[n]/m/	23	-0.41	.688	-0.17	0.15
	PLOSIVES: [d]/b/ -[b]/d/	23	2.78	.011* ²	-0.15	0.18
<i>Manner Contrasts</i>	CORONALS: [n]/d/ -[d]/n/	23	-3.79	.001*	1.35	0.48
	LABIALS: [b]/m/ -[m]/b/	23	-0.60	.552	0.44	0.38

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$;

²Note, the significance of [b]/d/, is a negativity, no positivity.

5.2.2.6 Experiment 6: Reaction-Time Experiment for Place & Manner Features

Table 19: Experiment 6: Statistical Analysis of Reaction Times for Place & Manner Features

Reaction Times [ms]							
source	level	df	F-value	p-value	mean	SEM	
<i>Place contrasts</i>	[n]/m/	15			669.81	11.73	
	[m]/n/	15			687.15	15.43	
	[d]/b/	15			614.82	11.15	
	[b]/d/	15			610.81	14.46	
	[d]/g/	15			633.42	8.97	
	[g]/d/	15			644.34	16.36	
<i>Manner contrasts</i>	[m]/b/	23			578.74	10.50	
	[b]/m/	23			575.81	10.88	
	[n]/d/	15			590.69	13.24	
	[d]/n/	15			558.61	12.29	
	[n]/z/	15			549.84	12.55	
	[z]/n/	15			515.67	14.34	
<i>Manner + Place contrasts</i>	[n]/g/	15			570.57	13.72	
	[g]/n/	15			569.57	16.60	
source	level	df	F-value	p-value	mean	SEM	
<i>Contrast</i>		6	17.00	.000*			
<i>Condition</i>		1	0.00	.986			
<i>Contrast x Condition</i>		6	2.85	.013*			
Place contrast							
<i>Feature-type</i>							
[LABIAL] ~ [CORONAL]	NASALS: [m]/n/ -[n]/m/	15	2.23	.156	678.48	12.42	
[LABIAL] ~ [CORONAL]	PLOSIVES: [d]/b/ -[b]/d/	15	0.11	.750	612.81	11.33	
[DORSAL] ~ [CORONAL]	PLOSIVES: [d]/g/ -[g]/d/	15	0.55	.472	638.88	10.93	
Manner contrast							
<i>Feature-type</i>							
[PLOSIVE] ~ [NASAL]	LABIALS: [b]/m/ -[m]/b/	23	0.08	.779	577.27	9.36	
[PLOSIVE] ~ [NASAL]	CORONALS: [n]/d/ -[d]/n/	15	5.70	.031*	574.65	10.87	
[STRIDENT] ~ [NASAL]	CORONALS: [n]/z/ -[z]/n/	15	11.48	.004*	532.76	12.50	
Manner + Place contrast							
<i>Feature-type</i>							
[PLOSIVE, DORSAL] ~ [NASALS, CORONAL]	[n]/g/ -[g]/n/	15	0.01	.931	570.07	14.11	

* $p < .05$; SEM = Standard Error of Mean; d.f. = degrees of freedom; # $p < .075$

6 References

- Aaltonen, O., Niemi, P., Nyrke, T. & Tuhkanen, M., (1987). 'Event-related brain potentials and the perception of a phonetic continuum'. *Biological Psychology*, 24 (3):197–207.
- Aaltonen, O., Eerola, O., Hellström, Å., Uusipaikka, E. & Lang, A., (1997). 'Perceptual magnet effect in the light of behavioral and psychophysiological data'. *The Journal of the Acoustical Society of America*, 101:1090.
- Aaltonen, O., Tuomainen, J., Laine, M. & Niemi, P., (1993). 'Cortical differences in tonal versus vowel processing as revealed by an ERP component called mismatch negativity (MMN)'. *Brain and Language*, 44:139-139.
- Alain, C., Woods, D. & Knight, R., (1998). 'A distributed cortical network for auditory sensory memory in humans'. *Brain Research*, 812 (1-2):23-37.
- Alho, K., (1995). 'Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes'. *Ear and Hearing*, 16 (1):38-50.
- Alho, K., Paavilainen, P., Reinikainen, K., Sams, M. & Näätänen, R., (1986). 'Separability of different negative components of the event-related potential associated with auditory stimulus processing'. *Psychophysiology*, 23 (6):613-623.
- Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K. & Näätänen, R., (1990). 'Event-related brain potential of human newborns to pitch change of an acoustic stimulus'. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 77 (2):151-155.
- Alho, K., Woods, D., Algazi, A., Knight, R. & Näätänen, R., (1994). 'Lesions of frontal cortex diminish the auditory mismatch negativity'. *Electroencephalography and Clinical Neurophysiology*, 91 (5):353-362.
- Allison, T., Wood, C. & McCarthy, G., (1986). 'The central nervous system'. In: M. G. H. Coles, E. Donchin, & S. W. Porges (Eds.) *Psychophysiology: Systems, Processes, and Applications* (pp. 5-25). Guilford, New York.
- Archangeli, D., (2008). 'Aspects of underspecification theory'. *Phonology*, 5 (02):183-207.
- Attias, J. & Pratt, H., (1992). 'Auditory event related potentials during lexical categorization in the oddball paradigm'. *Brain and Language*, 43 (2):230-239.
- Aulanko, R., Hari, R., Lounasmaa, O., Näätänen, R. & Sams, M., (1993). 'Phonetic invariance in the human auditory cortex'. *NeuroReport-Oxford-*, 4:1356-1356.
- Baayen, R., Piepenbrock, R. & Gulikers, L., (1995). 'The CELEX lexical database (release 2) [cd-rom]'. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].
- Baldeweg, T., Klugman, A., Gruzelier, J. & Hirsch, S., (2004). 'Mismatch negativity potentials and cognitive impairment in schizophrenia'. *Schizophrenia Research*, 69 (2-3):203-217.
- Baldeweg, T., Richardson, A., Watkins, S., Foale, C. & Gruzelier, J., (1999). 'Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials'. *Annals of Neurology*, 45 (4):495-503.

- Barceló, F., Escera, C., Corral, M.J., & Periáñez, J.A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *Journal of Cognitive Neuroscience*, 18, 17324–1738.
- Barlow, J.S. (1993). *The Electroencephalogram*. Cambridge: The MIT Press.
- Berger, H. (1929). Über das Elektroenzephalogramm des Menschen. *Archiv Psychiatrischer Nervenkrankheiten*, 87:527-570.
- Berti, S., & Schröger, E., (2001). A comparison of auditory and visual distraction effects: behavioral and event-related indices. *Cognitive Brain Research* 10, 265–273
- Bertrand, O., Perrin, F. & Pernier, J., (1991). 'Evidence for a tonotopic organization of the auditory cortex observed with auditory evoked potentials'. *Acta Oto-Laryngologica*, 111 (S491):116-123.
- Birbaumer, N. & Elbert, T., (1988). 'P3: By-product of a by-product'. *Behavioral and Brain Sciences*, 11:375-377.
- Birbaumer, N., Elbert, T., Canavan, A. & Rockstroh, B., (1990). 'Slow potentials of the cerebral cortex and behavior'. *Physiological Reviews*, 70 (1):1-41.
- Boersma, P. & Weenink, D., (2007). 'PRAAT: Doing Phonetics by Computer (ver. 4.6.38)'. Amsterdam: Institut for Phonetic Sciences.
- Bonte, M.L., Mitterer, H., Zellagui, N., Poelmans, H. & Blomert, L., (2005). 'Auditory cortical tuning to statistical regularities in phonology'. *Clinical Neurophysiology*, 116:2765-2774.
- Bonte, M., Poelmans, H. & Blomert, L., (2007). 'Deviant neurophysiological responses to phonological regularities in speech in dyslexic children'. *Neuropsychologia*, 45 (7):1427-1437.
- Bradlow, A., Kraus, N., Nicol, T., McGee, T., Cunningham, J., Zecker, S. & Carrell, T., (1999). 'Effects of lengthened formant transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children'. *The Journal of the Acoustical Society of America*, 106:2086.
- Braitenberg, W. (1977). *On the Texture of Brains*. Springer.
- Bybee, J., (2001). *Phonology and Language Use*. Cambridge, UK: Cambridge University Press.
- Bybee, J. & Schreibman, J., (1999). 'The effect of usage on degrees of constituency: the reduction of don't in English'. *Linguistics*, 37 (4):575-596.
- Caramazza, A., (1996). 'Neuropsychology: The brain's dictionary'. *Nature*, 380 (6574):485-486.
- Caton, R. (1875). The electric currents of the brain. *British Medical Journal*, 2: 278.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K. & Näätänen, R., (1998). 'Development of language-specific phoneme representations in the infant brain'. *Nature Neuroscience*, 1 (5):351-353.

- Chomsky, N. and Halle, M., 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Clements, G.N., (1989). '*A unified set of features for consonants and vowels*'. Cornell University, Ithaca, New York.
- Clements, G. N., & Hume, E. V. (1995). The internal organization of speech sounds. In: Goldsmith, J.A., (Ed.), *The Handbook of Phonological Theory* (pp. 245-306). Oxford: Blackwell.
- Coles, M., (1989). 'Modern mind-brain reading: Psychophysiology, physiology, and cognition'. *Psychophysiology*, 26 (3):251-269.
- Coles, M., Smid, H., Scheffers, M., & Otten, L.J. (1995). Mental chronometry and the study of human information processing. In: M.D. Rugg & M. Coles (Eds.) *Electrophysiology of mind event-related brain potentials and cognition* (pp. 87–131). Oxford, UK: Oxford University Press.
- Cole, R., (1973). 'Different memory functions for consonants and vowels'. *Cognitive Psychology*, 4 (1):39-54.
- Colin, C., Radeau, M., Soquet, A. & Deltenre, P., (2004). 'Generalization of the generation of an MMN by illusory McGurk percepts: voiceless consonants'. *Clinical Neurophysiology*, 115 (9):1989-2000.
- Combs, L. & Polich, J., (2006). 'P3a from auditory white noise stimuli'. *Clinical Neurophysiology*, 117 (5):1106-1112.
- Connine, C.M., Blasko, D. & Titone, D., (1993). 'Do the beginnings of spoken words have a special status in auditory word recognition?'. *Journal of Memory and Language*, 32 (2):193-210.
- Cooper, R., Osselton, J. & Shaw, J., (1984). '*Elektroenzephalographie: Technik und Methoden*'. Aufl. Fischer, Stuttgart.
- Cornell, S.A., Lahiri, A. & Eulitz, C., (2008). 'What you encode is not necessarily what you store: An MMN study of phonological processing'. *Supplement of the Journal of Cognitive Neuroscience* 290.
- Cornell, A.S., Lahiri, A. & Eulitz, C., (2009). 'MMN to phonetic contrasts along different feature dimensions.'. *Frontiers in Human Neuroscience. Conference Abstract: MMN 09 Fifth Conference on Mismatch Negativity (MMN) and its Clinical and Scientific Applications in Budapest*, April. doi: s10.3389/conf.neuro.09.2009.05.079
- Cornell, A.S., Lahiri, A., & Eulitz, C. (submitted). 'Inequality in Manner-Of-Articulation Contrasts In Speech Perception: Evidence From Event-Related Brain Potentials'. *Journal of Experimental Psychology: Human Perception and Performance*.
- Courchesne, E., Hillyard, S. & Galambos, R., (1975). 'Stimulus novelty, task relevance and the visual evoked potential in man'. *Electroencephalography and Clinical Neurophysiology*, 39 (2):131-143.

- Cowan, N., Winkler, I., Teder, W. & Näätänen, R., (1993). 'Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP)'. *Journal of Experimental Psychology Learning, Memory and Cognition*, 19 (4):909-921.
- Cowan, N., (1984). 'On short and long auditory stores'. *Psychological Bulletin.*, 96 (2):341-370.
- Crowder, R., (1971). 'The sound of vowels and consonants in immediate memory'. *Journal of Verbal Learning and Verbal Behavior*. Vol, 10 (6):587-596.
- Cutler, E. & Norris, D., (1979). 'Monitoring sentence comprehension'. In: W.E. Cooper & E.C.T. Walker (Eds.), *Sentence Processing: Psycholinguistic Studies presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.
- Dalebout, S. & Stack, J., (1999). 'Mismatch negativity to acoustic differences not differentiated behaviorally'. *Journal of the American Academy of Audiology*, 10 (7):388.
- Dalebout, S. & Fox, L., (2000). 'Identification of the mismatch negativity in the responses of individual listeners'. *Journal of the American Academy of Audiology*, 11 (1):12.
- Dehaene-Lambertz, G., (1997). 'Electrophysiological correlates of categorical phoneme perception in adults'. *NeuroReport*, 8:919-924.
- Dehaene-Lambertz, G., Dupoux, E. and Gout, A., (2000). 'Electrophysiological correlates of phonological processing: a cross-linguistic study'. *Journal of Cognitive Neuroscience*, 12 (4):635-647.
- Dehaene-Lambertz, G. & Gliga, T., (2004). 'Common neural basis for phoneme processing in infants and adults'. *Journal of Cognitive Neuroscience*, 16 (8):1375-1387.
- Deouell, L., (2007). 'The frontal generator of the mismatch negativity revisited'. *Journal of Psychophysiology*, 21 (3/4):188.
- Diesch, E., Biermann, S. & Luce, T., (1998). 'The magnetic mismatch field elicited by words and phonological non-words'. *NeuroReport*, 9 (3):455.
- Diesch, E. & Luce, T., (1997). 'Magnetic mismatch fields elicited by vowels and consonants'. *Experimental Brain Research*, 116 (1):139-152.
- Dimoska, A., Johnstone, S. & Barry, R., (2006). 'The auditory-evoked N2 and P3 components in the stop-signal task: Indices of inhibition, response-conflict or error-detection?' *Brain and Cognition*, 62 (2):98-112.
- Donchin, E., (1981). 'Surprise!... surprise'. *Psychophysiology*, 18 (5):493-513.
- Donchin, E. & Coles, M., (1988). 'Is the P300 component a manifestation of context updating?' *Behavioral and Brain Sciences*, 11 (3):357-427.
- Ellis, A. & Young, A., (1996). *Human Cognitive Neuropsychology: A Textbook with Readings*: Psychology Press (UK).
- Elman, J., (1990). 'Finding structure in time'. *Cognitive Science*, 14, 179-211.

- Endrass, T., Mohr, B. & Pulvermüller, F., (2004). 'Enhanced mismatch negativity brain response after binaural word presentation'. *European Journal of Neuroscience*, 19 (6):1653-1660.
- Escera, C., (2007). 'The Mismatch Negativity 30 Years Later: How Far Have We Come?' *Journal of Psychophysiology*, 21 (3/4):129.
- Escera, C. & Corral, M., (2007). 'Role of mismatch negativity and novelty-P3 in involuntary auditory attention'. *Journal of Psychophysiology*, 21 (3/4):251.
- Escera, C., Alho, K., Schröger, E. & Winkler, I., (2000). 'Involuntary attention and distractibility as evaluated with event-related brain potentials'. *Audiology and Neurotology*, 5 (3-4):151-166.
- Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, 10, 590-604.
- Escera, C., Yago, E., & Alho, K. (2001). Electrical responses reveal the temporal dynamics of brain events during involuntary attention switching. *European Journal of Neuroscience*, 14, 877–883.
- Eulitz, C. & Lahiri, A., (2004). 'Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition'. *Journal of Cognitive Neuroscience*, 16 (4):577-583.
- Fabiani, M., Friedman, D. & Cheng, J., (1998). 'Individual differences in P3 scalp distribution in older adults, and their relationship to frontal lobe function'. *Psychophysiology*, 35 (06):698-708.
- Felder, V., (2006). 'Examining featural underspecification of Tongue Height in German mid vowels: an EEG study'. *University of Konstanz: Diplomarbeit*.
- Felder, V., Friedrich, C.K., Lahiri, A. & Eulitz, C., (2008). 'Unterspezifikation von Vokalhöhe im Mentalen Lexikon - Ereigniskorrelierte Potentiale versus Verhaltensdaten.' In V. Felder (Chair): *Der Teufel sitzt im Detail - Wieviel phonologische Spezifikation braucht das Lexikon?* Symposium conducted at the 50.Tagung experimentell arbeitender Psychologen, Marburg, Germany.
- Fischer, C., Morlet, D., Bouchet, P., Luaute, J., Jourdan, C. & Salord, F., (1999). 'Mismatch negativity and late auditory evoked potentials in comatose patients'. *Clinical Neurophysiology*, 110 (9):1601-1610.
- Fitzpatrick, J. & Wheeldon, L.R., (2000). 'Phonology and Phonetics in Psycholinguistic Models of Speech Perception'. In: N. Burton-Roberts, P. Carr, & G. Docherty (Eds). *Phonological Knowledge: Conceptual and Empirical Issues* (pp. 131-160). Oxford: Oxford University Press.
- Frauenfelder, U. & Tyler, L., (1987). 'The process of spoken word recognition: An introduction'. *Cognition*, 25 (1-2):1-20.
- Friedman, D. & Simpson, G., (1994). 'ERP amplitude and scalp distribution to target and novel events: Effects of temporal order in young, middle-aged and older adults'. *Cognitive Brain Research*, 2 (1):49-63.

- Friedman, D., Cycowicz, Y. & Gaeta, H., (2001). 'The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty'. *Neuroscience and Biobehavioral Reviews*, 25 (4):355-373.
- Friedrich, C.K., Eulitz, C. & Lahiri, A., (2006). 'Not every pseudoword disrupts word recognition: an ERP study'. *Behavioral and Brain Function*, 2:36.
- Friedrich, C.K., Lahiri, A. & Eulitz, C., (2008). 'Neurophysiological evidence for underspecified lexical representations: asymmetries with word initial variations.' *Journal of Experimental Psychology: Human Perception and Performance.*, 34 (6):1545-1559.
- Gaeta, H., Friedman, D., Ritter, W. & Cheng, J., (2001). 'An event-related potential evaluation of involuntary attentional shifts in young and older adults'. *Psychology and Aging*, 16 (1):55-68.
- Gaskell, M., (2003). 'Modelling regressive and progressive effects of assimilation in speech perception'. *Journal of Phonetics*, 31 (3-4):447-463.
- Gaskell, M.G., Hare, M. & Marslen-Wilson, W.D., (1995). 'A connectionist model of phonological representation in speech perception'. *Cognitive Science: A Multidisciplinary Journal of Artificial Intelligence, Linguistics, Neuroscience, Philosophy, Psychology*, 19 (4):407-439.
- Gaskell, M.G. & Marslen-Wilson, W.D., (1996). 'Phonological variation and inference in lexical access'. *Journal of Experimental Psychology: Human Perception and Performance*, 22 (1):144-158.
- Gaskell, M.G. & Marslen-Wilson, W.D., (1997). 'Integrating form and meaning: A distributed model of speech perception'. *Language and Cognitive Processes*, 12 (5-6):613-656.
- Gaskell, M.G. & Marslen-Wilson, W.D., (1998). 'Mechanisms of phonological inference in speech perception'. *Journal of Experimental Psychology-Human Perception and Performance*, 24 (2):380-396.
- Gaskell, M.G. & Marslen-Wilson, W.D., (1999). 'Ambiguity, competition, and blending in spoken word recognition'. *Cognitive Science*, 23 (4):439-462.
- Gaskell, M.G. & Marslen-Wilson, W.D., (2001). 'Lexical ambiguity resolution and spoken word recognition: Bridging the gap'. *Journal of Memory and Language*, 44 (3):325-349.
- Gaskell, M. & Snoeren, N., (2008). 'The impact of strong assimilation on the perception of connected speech'. *Journal of Experimental Psychology: Human Perception and Performance*, 34:1632-1647.
- Gazzaniga, M., Ivry, R. & Mangun, G. (2002). *Cognitive Neuroscience: The Biology of the Mind* (2nd ed.). WW Norton & Company.
- Ghini, M., (2001). *Asymmetries in the Phonology of Miogliola*. Berlin: Mouton de Gruyter.
- Giard, M.H., Perrin, F., Echallier, J.F., Thévenet, M., Froment, J.C., & Pernier, J. (1994). Dissociation of temporal and frontal components in the auditory N1 wave: A scalp

- current density and dipole model analysis. *Electroencephalography and Clinical Neurophysiology*, 92, 238–252.
- Giard, M., Perrin, F., Pernier, J. & Bouchet, P., (1990). 'Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study'. *Psychophysiology*, 27 (6):627-640.
- Goldinger, S.D., (1998). 'Echoes of echoes? An episodic theory of lexical access'. *Psychological Review*, 105 (2):251-279.
- Gomes, H., Ritter, W. & Vaughan Jr, H., (1995). 'The nature of preattentive storage in the auditory system'. *Journal of Cognitive Neuroscience*, 7 (1):81-94.
- Gow, D.W., Jr., (2002). 'Does English coronal place assimilation create lexical ambiguity?' *Journal of Experimental Psychology: Human Perception and Performance*, 28 (1):163-179.
- Gow, D.W., Jr., (2003). 'Feature parsing: feature cue mapping in spoken word recognition'. *Perception and Psychophysics*, 65 (4):575-590.
- Grosjean, F., (1980). 'Spoken word recognition processes and the gating paradigm'. *Perception and Psychophysics*, 28 (4):267-283.
- Gumnior, H., Zwitserlood, P. & Bölte, J., (2005). 'Assimilation in existing and novel German compounds'. *Language and Cognitive Processes*, 20 (3):465-488.
- Hacquard, V., Walter, M. & Marantz, A., (2007). 'The effects of inventory on vowel perception in French and Spanish: An MEG study'. *Brain and Language*, 100 (3):295-300.
- Hagoort, P., (2008). 'The fractionation of spoken language understanding by measuring electrical and magnetic brain signals'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363 (1493):1055-1069.
- Halle, M. (1995). Feature geometry and feature spreading. *Linguistic Inquiry*, 26, 1-46.
- Halle, M. (1992). Phonological features. In: Bright, W., (Ed.), *International encyclopedia of linguistics* (Vol. 3, pp. 207–212). Oxford: Oxford University Press.
- Halle, M., Hughes, G. & Radley, J.A. (1957): Acoustic properties of stop consonants. *Journal of the Acoustical Society of America (JASA)*, 29(1), 107-116.
- Halle, M., Vaux, B. & Wolfe, A., (2000). 'On feature spreading and the representation of place of articulation'. *Linguistic Inquiry*, 31:387-444.
- Halle, M. & Stevens, K., (1991). 'Knowledge of language and the sounds of speech'. In: J. Sundberg, L. Nord, & R. Carlson, (Eds.) *Music, Language, Speech and Brain* (pp. 1-19), McMillan, London.
- Hari, R., Kaila, K., Katila, T., Tuomisto, T. & Varpula, T., (1982). 'Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: implications for their neural generation'. *Electroencephalography and Clinical Neurophysiology*, 54 (5):561-569.
- Hickok, G. & Poeppel, D., (2000). 'Towards a functional neuroanatomy of speech

- perception'. *Trends in Cognitive Science*, 4 (4):131-138.
- Hickok, G. & Poeppel, D., (2007). 'Opinion – The cortical organization of speech processing'. *Nature Reviews Neuroscience*, 8 (5):393-402.
- Holm, A., Ranta-aho, P., Sallinen, M., Karjalainen, P. & Müller, K., (2006). 'Relationship of P300 single-trial responses with reaction time and preceding stimulus sequence'. *International Journal of Psychophysiology*, 61 (2):244-252.
- Horváth, J., Winkler, I. & Bendixen, A., (2008). 'Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction?' *Biological Psychology*, 79 (2):139-147.
- Huutilainen, M., Ilmoniemi, R., Lavikainen, J., Tiitinen, H., Alho, K., Sinkkonen, J., Knuutila, J., & Näätänen, R. (1993). Interaction between representations of different features of auditory sensory memory. *NeuroReport*, 4, 1279–1281.
- Imada, T., Fukuda, K., Kawakatsu, M., Mashiko, T., Okada, K., Hayashi, M., Aihara, K. & Kotani, M., (1995). 'Mismatch fields evoked by a rhythm passage'. *Biomagnetism: Fundamental Research and Clinical Applications*. Elsevier, Amsterdam:249–252.
- Imada, T., Hari, R., Loveless, N. & McEvoy, L., (1993). 'Determinants of the auditory mismatch response'. *Electroencephalography and Clinical Neurophysiology*, 87 (3):144-153.
- Jackendoff, R., (2003). *Foundations of language: Brain, Meaning, Grammar, Evolution*: Oxford University Press, USA.
- Jacobsen, T., Horváth, J., Schröger, E., Lattner, S., Widmann, A. & Winkler, I., (2004). 'Pre-attentive auditory processing of lexicality'. *Brain and Language*, 88:54-67.
- Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to Speech Analysis. The Distinctive Features and Their Correlates*. Cambridge, MA: MIT, Acoustics Laboratory, Technical Report No. 13.
- Johnson, R., (1993). 'On the neural generators of the P300 component of the event-related potential'. *Psychophysiology*, 30:90-90.
- Johnson, K., (1997). 'Speech perception without speaker normalization: An exemplar model'. In: K. Johnson & J.W. Mullenix (Eds). *Talker Variability in Speech Processing*. San Diego: Academic Press.
- Katayama, J. & Polich, J., (1998). 'Stimulus context determines P3a and P3b'. *Psychophysiology*, 35 (01):23-33.
- King, C., McGee, T., Rubel, E., Nicol, T. & Kraus, N., (1995). 'Acoustic features and acoustic change are represented by different central pathways'. *Hearing Research*, 85 (1-2):45-52.
- Klatt, D.H., (1989). 'Review of selected models of speech perception'. In: W. Marslen-Wilson (Ed). *Lexical Representation and Process*. Cambridge, MA: MIT Press, 169-226.
- Knight, R., (1984). 'Decreased response to novel stimuli after prefrontal lesions in man'. *Electroencephalography and Clinical Neurophysiology*, 59 (1):9-20.

- Knight, R., (1996). 'Contribution of human hippocampal region to novelty detection'. *Nature*, 383 (6597):256-259.
- Knight, R., (1997). 'Distributed cortical network for visual attention'. *Journal of Cognitive Neuroscience*, 9 (1):75-91.
- Knight, R., Scabini, D., Woods, D. & Clayworth, C., (1989). 'Contributions of temporal-parietal junction to the human auditory P3'. *Brain Research*, 502 (1):109-116.
- Kohler, K.J., (1995). *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt Verlag.
- Korpilahti, P., Krause, C., Holopainen, I. & Lang, A., (2001). 'Early and late mismatch negativity elicited by words and speech-like stimuli in children'. *Brain and Language*, 76 (3):332-339.
- Kraus, N. & Cheour, M., (2000). 'Speech sound representation in the brain'. *Audiology and Neuro-Otology*, 5 (3-4):140-150.
- Kraus, N., McGee, T., Carrell, T., Sharma, A., Micco, A. & Nicol, T., (1993a). 'Speech-evoked cortical potentials in children'. *Journal of the American Academy of Audiology*, 4 (4):238-248.
- Kraus, N., McGee, T., Sharma, A., Carrell, T. & Nicol, T., (1992). 'Mismatch negativity event-related potential elicited by speech stimuli'. *Ear and Hearing*, 13 (3):158-164.
- Kruschke, J., (1992). 'ALCOVE: An exemplar-based connectionist model of category learning'. *Psychological Review*, 99 (1):22-44.
- Kuhl, P., (1991). 'Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not'. *Perception and Psychophysics*, 50 (2):93-107.
- Kujala, A., Alho, K., Valle, S., Sivonen, P., Ilmoniemi, R., Alku, P. & Näätänen, R., (2002). 'Context modulates processing of speech sounds in the right auditory cortex of human subjects'. *Neuroscience Letters*, 331 (2):91-9
- Kujala, T., Karma, K., Ceponiene, R., Belitz, S., Turkkila, P., Tervaniemi, M. & Näätänen, R., (2001b). 'Plastic neural changes and reading improvement caused by audiovisual training in reading-impaired children'. *Proceedings of the National Academy of Sciences*, 98 (18):10509.
- Kurowski, K. & Blumstein, S., (1984). 'Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants'. *The Journal of the Acoustical Society of America*, 76:383.
- Kutas, M. & Federmeier, K.D., (2000). 'Electrophysiology reveals semantic memory use in language comprehension'. *Trends in Cognitive Sciences*, 4 (12):463-472.
- Kutas, M. & Hillyard, S.A., (1980). 'Reading senseless sentences: brain potential reflect semantic incongruity'. *Science*, 207 (4427):203-205.
- Lahiri, A. & Reetz, H., (2002). 'Underspecified recognition'. In: Gussenhoven, C. and Warner, N. (Eds). *Laboratory Phonology VII*. Berlin: Mouton de Gruyter, 637-677.

- Lahiri, A. & Marslen-Wilson, W., (1991). 'The mental representation of lexical form: A phonological approach to the recognition lexicon'. *Cognition*, 38:245-294.
- Lahiri, A. & Marslen-Wilson, W., (1992). 'Lexical processing and phonological representation'. In: Docherty, G. and Ladd, R. (eds). *Papers in Laboratory Phonology, II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, 229-254.
- Lahiri, A. & Reetz, H., (2010). 'Distinctive Features: Phonological underspecification in representation and processing'. *Journal of Phonetics*, 38, 44-59.
- Lahiri, A., (1999). 'Speech recognition with phonological features'. *The XIVth International Congress of Phonetic Sciences*. San Francisco.
- Lahiri, A., (2007). 'Non-equivalence between phonology and phonetics'. *XVI the International Conference of Phonetic Sciences*. Saarbrücken.
- Lahiri, A., (2000). 'Phonology: structure, representation, and process'. In: Wheeldon, L. (Ed). *Aspects of Language Production Studies in Cognition Series* (pp. 165-225). Philadelphia, PA: Psychology Press,
- Lahiri, A. & Evers, V., (1991). 'Palatalization and coronality'. In: Paradis, C. and Prunet, J.-F. (Eds). *Phonetics and Phonology Volume II: The Special Status of Coronals: Internal and External Evidence* (pp. 79-100). San Diego: Academic Press.
- Lagopoulos, J., Gordon, E., Barhamali, H., Lim, C., Li, W., Clouston, P. & Morris, J., (1998). 'Dysfunctions of automatic (P300a) and controlled (P300b) processing in Parkinson's disease'. *Neurological Research*, 20 (1):5-10.
- Lang, A.H., Nyrke, T., Ek, M., Aaltonen, O., Raimo, I., & Näätänen, R., (1990). Pitch discrimination performance and auditory event-related potentials. In: C.M.H. Brunia, A.W.K. Gaillard, A. Kok (Eds.) *Psychophysiological Brain Research, Vol I* (pp. 294-298). Tilburg University Press, Tilburg.
- Lau, E., Phillips, C. & Poeppel, D., (2008). 'A cortical network for semantics: (de) constructing the N400'. *Nature Reviews Neuroscience*, 9 (12):920-933.
- Lesser, R. & Milroy, L., (1993). *Linguistics and Aphasia: Psycholinguistic and Pragmatic Aspects of Intervention*: Longman Pub Group.
- Linden, D., (2005). 'The p300: where in the brain is it produced and what does it tell us?' *The Neuroscientist*, 11 (6):563-576.
- Luck, S., (2005). *An Introduction to the Event-related Potential Technique*: MIT Press Cambridge, MA.
- Lyytinen, H., Blomberg, A. & Näätänen, R., (1992). 'Event-related potentials and autonomic responses to a change in unattended auditory stimuli'. *Psychophysiology*, 29:523-523.
- Maiste, A.C., Wiens, A.S., Hunt, M.J., Scherg, M. & Picton, T.W., (1995). 'Event-related potentials and the categorical perception of speech sounds'. *Ear and Hearing*, 16 (1):68-90.

- Marslen-Wilson, W. & Tyler, L.K., (1980). 'The temporal structure of spoken language understanding'. *Cognition*, 8:1-71.
- Marslen-Wilson, W. & Welsh, A., (1978). 'Processing interactions and lexical access during word recognition in continuous speech'. *Cognitive Psychology*, 10:29-63.
- May, P., Tiitinen, H., Ilmoniemi, R., Nyman, G., Taylor, J. & Näätänen, R., (1999). 'Frequency change detection in human auditory cortex'. *Journal of Computational Neuroscience*, 6 (2):99-120.
- Mäkelä, J., Salmelin, R., Hari, R., & Hukkanen, S., *personal communication*, May 2, 1994
- McCarthy, G. & Donchin, E., (1981). 'A metric for thought: A comparison of P300 latency and reaction time'. *Science*, 211 (4477):77-80.
- McClelland, J.L. & Elman, J.L., (1986). 'The TRACE model of speech perception'. *Cognitive Psychology*, 18:1-86.
- McQueen, J., (2004). 'Speech perception. In: K. Lamberts, & R.L. Goldstone (Eds.) *The Handbook of Cognition* (pp. 255-275). Sage Publications Ltd. The Cromwell Press Ltd., Trowbridge, Wiltshire.
- Menning, H., Zwitserlood, P., Schöning, S., Hihn, H., Bölte, J., Dobel, C., Mathiak, K. & Lutkenhoner, B., (2005). 'Pre-attentive detection of syntactic and semantic errors'. *NeuroReport*, 16 (1):77-80.
- Metcalfe, J., (1993). 'Monitoring and gain control in an episodic memory model: Relation to the P300 event-related potential'. *Theories of Memory*: 327-353.
- Mitterer, H. (2003). Understanding Gardem bench: Studies on the perception of assimilated word forms [*Dissertation*]. Maastricht, The Netherlands: Universiteit Maastricht.
- Mitterer, H. & Blomert, L., (2003). 'Coping with phonological assimilation in speech perception: Evidence for early compensation'. *Perception and Psychophysics*, 65 (6):956-969.
- Müller, B., Jüptner, M., Jentzen, W. & Müller, S., (2002). 'Cortical activation to auditory mismatch elicited by frequency deviant and complex novel sounds: a PET study'. *NeuroImage*, 17 (1):231-239.
- Näätänen, R., (1990). 'The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function'. *Behavioral and Brain Sciences*, 13 (2):201-288.
- Näätänen, R., (1992). '*Attention and Brain Function*'. Erlbaum, Hillsdale.
- Näätänen, R., (1995). 'The mismatch negativity: a powerful tool for cognitive neuroscience'. *Ear and Hearing*, 16 (1):6-18.
- Näätänen, R., (2001). 'The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm)'. *Psychophysiology*, 38:1-21.
- Näätänen, R. & Alho, K., (1997). 'Mismatch negativity-the measure for central sound representation accuracy'. *Audiology and Neuro-Otology*, 2 (5):341-353.

- Näätänen, R. & Alho, K., (1995). 'Mismatch negativity-a unique measure of sensory processing in audition'. *International Journal of Neuroscience*, 80 (1-4):317-337.
- Näätänen, R., Ilmoniemi, R. & Alho, K., (1994). 'Magnetoencephalography in studies of human cognitive brain function'. *Trends in Neurosciences*, 17 (9):389-395.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J. & Alho, K., (1997). 'Language-specific phoneme representations revealed by electric and magnetic brain responses'. *Nature*, 385 (6615):432-434.
- Näätänen, R. & Michie, P., (1979). 'Early selective attention effects on the evoked potential: A critical review and reinterpretation'. *Biological Psychology*, 8 (2):81-136.
- Näätänen, R., Gaillard, A. & Mäntysalo, S., (1978). 'Early selective-attention effect on evoked potential reinterpreted'. *Acta Psychologica*, 42 (4):313-329.
- Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K. & Sams, M., (1987). 'The mismatch negativity to intensity changes in an auditory stimulus sequence'. *Electroencephalography and Clinical Neurophysiology*, 40:125-131.
- Näätänen, R., Paavilainen, P. & Reinikainen, K., (1989). 'Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man?' *Neuroscience Letters*, 107 (1-3):347-352.
- Näätänen, R., Paavilainen, P., Rinne, T. & Alho, K., (2007). 'The mismatch negativity (MMN) in basic research of central auditory processing: A review'. *Clinical Neurophysiology*, 118 (12):2544-2590.
- Näätänen, R. & Picton, T., (1987). 'The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure'. *Psychophysiology*, 24 (4):375-425.
- Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M. & Paavilainen, P., (1993). 'Development of a memory trace for a complex sound in the human brain'. *NeuroReport - Oxford* -, 4:503-503.
- Näätänen, R. & Winkler, I., (1999). 'The concept of auditory stimulus representation in cognitive neuroscience'. *Psychological Bulletin*, 125 (6):826-859.
- Nieuwenhuis, S., Aston-Jones, G. & Cohen, J., (2005). 'Decision making, the P3, and the locus coeruleus-norepinephrine system'. *Psychological Bulletin*, 131 (4):510-532.
- Nolan, F., (1992). 'The descriptive role of segments: evidence from assimilation'. *Papers in Laboratory Phonology*, 2:261-280.
- Norris, D., (1994). 'Shortlist: A connectionist model of continuous speech recognition'. *Cognition*, 52:189-234.
- Norris, D., (1990). 'A dynamic-net model of human speech recognition'. *Cognitive Models of Speech Processing*:87-104.
- Norris, D. & McQueen, J., (2008). 'Shortlist B: A Bayesian model of continuous speech recognition'. *Psychological Review*, 115 (2):357-395.

- Obleser, J., Elbert, T., Lahiri, A. & Eulitz, C., (2003). 'Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies'. *Cognitive Brain Research*, 15 (3):207-213.
- Obleser, J., Eulitz, C. & Lahiri, A., (2004). 'Magnetic brain response mirrors extraction of phonological features from spoken vowels'. *Journal of Cognitive Neuroscience*, 16 (1):31-39.
- Öhman A. (1979). The orienting response, attention and learning: an information processing perspective. In: H.D. Kimmel, E.H. van Olst, & J.F. Orlebeke (Eds.). *The Orienting Response in Humans* (p. 443-471). Hillsdale, NJ: Laurence Erlbaum Associates.
- Oldfield, R.C., (1971). 'The assessment and analysis of handedness: The Edinburgh Inventory'. *Neuropsychologia*, 9:97-113.
- O'Leary, D., Andreasen, N., Hurtig, R., Hichwa, R., Watkins, G., Boles Ponto, L., Rogers, M. & Kirchner, P., (1996). 'A positron emission tomography study of binaurally and dichotically presented stimuli: Effects of level of language and directed attention'. *Brain and Language*, 53 (1):20-39.
- Opitz, B., Mecklinger, A., von Cramon, D. & Kruggel, F., (1999). 'Combining electrophysiological and hemodynamic measures of the auditory oddball'. *Psychophysiology*, 36 (01):142-147.
- Ottaviani, F., Girolamo, S., Briglia, G., Rossi, G., Giuda, D. & Nardo, W., (1997). 'Tonotopic organization of human auditory cortex analyzed by SPET'. *International Journal of Audiology*, 36 (5):241-248.
- Pallier, C., Colomé, A. & Sebastian-Galles, N., (2001). 'The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries'. *Psychological Science*, 12 (6):445-449.
- Paradis, C. & Prunet, J.-F., (1991). *Phonetics and Phonology Volume II: The Special Status of Coronals: Internal and External Evidence*. San Diego: Academic Press.
- Pekkonen, E., Jousmäki, V., Kononen, M., Reinikainen, K. & Partanen, J., (1994). 'Auditory sensory memory impairment in Alzheimer's disease: an event-related potential study'. *NeuroReport*, 5:2537-2540.
- Pettigrew, C., Murdoch, B., Chenery, H., & Kei, J., (2004). 'The relationship between the mismatch negativity (MMN) and psycholinguistic models of spoken word processing'. *Aphasiology*, 18 (1):3-28.
- Pettigrew, C.M., Murdoch, B.E., Ponton, C.W., Finnigan, S., Alku, P., Kei, J., Sockalingam, R. & Chenery, H.J., (2004a). 'Automatic auditory processing of English words as indexed by the mismatch negativity, using a multiple deviant paradigm'. *Ear and Hearing*, 25 (3):284-301.
- Pettigrew, C.M., Murdoch, B.M., Kei, J., Chenery, H.J., Sockalingam, R., Ponton, C.W., Finnigan, S. & Alku, P., (2004b). 'Processing of English words with fine acoustic contrasts and simple tones: a mismatch negativity study'. *Journal of the American Academy of Audiology*, 15 (1):47-66.

- Pisoni, D. & Luce, P., (1987). 'Acoustic-phonetic representations in word recognition'. *Cognition*, 25 (1-2):21–52.
- Phillips, C., (2001). 'Levels of representation in the electrophysiology of speech perception'. *Cognitive Science*, 25 (5):711-731.
- Phillips, C., Marantz, A., McGinnis, M., Pesetsky, D., Wexler, K., Yellin, A., Poeppel, D., Roberts, T. & Rowley, H., (1995). 'Brain mechanisms of speech perception: A preliminary report'. *MIT Working Papers in Linguistics*, 26:125–163.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M. & Roberts, T., (2000a). 'Auditory Cortex Accesses Phonological Categories: An MEG Mismatch Study'. *Journal of Cognitive Neuroscience*, 12 (6):1038-1055.
- Phillips, C., Pellathy, T. & Marantz, A., (2000b). 'Phonological feature representations in auditory cortex'. *Submitted Manuscript*.
- Picton, T., (1992). 'The P300 wave of the human event-related potential'. *Journal of Clinical Neurophysiology*, 9 (4):456.
- Picton, T.W., (1995). 'The neurophysiological evaluation of auditory discrimination'. *Ear and Hearing*, 16 (1):1-5.
- Picton, T.W., Alain, C., Otten, L., Ritter, W. & Achim, A., (2000). 'Mismatch negativity: Different water in the same river'. *Audiology & Neuro-Otology*, 5:111-139.
- Picton, T.W. & S.A. Hillyard (1988). Endogenous event-related potentials. In: T.W. Picton (Ed.). *Handbook of Electroencephalography and Clinical Neurophysiology Human Event-Related Potentials* (Vol. 3; pp. 361-426). Amsterdam: Elsevier.
- Poeppel, D., Idsardi, W.J. & van Wassenhove, V., (2008). 'Speech perception at the interface of neurobiology and linguistics'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363 (1493):1071-1086.
- Poeppel, D., (2003). 'The analysis of speech in different temporal integration windows: cerebral lateralization as' asymmetric sampling in time". *Speech Communication*, 41 (1):245-256.
- Polich, J., (2004). Neuropsychology of the P3a and P3b: A theoretical overview. In: N.C. Moore & K. Arikan (Eds.), *Brainwaves and Mind: Recent Developments* (pp. 15-29). Kjellberg Inc.: Wheaton, IL.
- Polich, J., (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology* 118, 2128–2148.
- Polich, J. & Comerchero, M., (2003). 'P3a from visual stimuli: typicality, task, and topography'. *Brain Topography*, 15 (3):141-152.
- Posner, M. & Petersen, S., (1990). 'The attention system of the human brain'. *Annual Review of Neuroscience*, 13 (1):25-42.
- Pulvermüller, F. & Assadollahi, R., (2007). 'Grammar or serial order?: Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity'. *Journal of Cognitive Neuroscience*, 19 (6):971-980.

- Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Illmoniemi, R.J. & Näätänen, R., (2001). 'Memory traces for words as revealed by the mismatch negativity'. *NeuroImage*, 14:607-616.
- Pulvermüller, F. & Shtyrov, Y., (2003). 'Automatic processing of grammar in the human brain as revealed by the mismatch negativity'. *NeuroImage*, 20 (1):159-172.
- Pulvermüller, F. & Shtyrov, Y., (2006a). 'Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes'. *Progress in Neurobiology*, 79:49-71.
- Pulvermüller, F., Shtyrov, Y., Illmoniemi, R.J. & Marslen-Wilson, W., (2006b). 'Tracking speech comprehension in space and time'. *NeuroImage*, 31:1297-1305.
- Pulvermüller, F., Shtyrov, Y. & Ilmoniemi, R., (2005). 'Brain signatures of meaning access in action word recognition'. *Journal of Cognitive Neuroscience*, 17 (6):884-892.
- Pulvermüller, F., Shtyrov, Y., Kujala, T. & Näätänen, R., (2004). 'Word-specific cortical activity as revealed by the mismatch negativity'. *Psychophysiology*, 41 (1):106-112.
- Rinne, T., Gratton, G., Fabiani, M., Cowan, N., Maclin, E., Stinard, A., Sinkkonen, J., Alho, K. & Näätänen, R., (1999b). 'RAPID COMMUNICATION Scalp-Recorded Optical Signals Make Sound Processing in the Auditory Cortex Visible?' *NeuroImage*, 10 (5):620-624.
- Ritter, W., Deacon, D., Gomes, H., Javitt, D.C. & Vaughan, H.G., Jr., (1995). 'The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review'. *Ear and Hearing*, 16 (1):52-67.
- Rivera-Gaxiola, M., Csibra, G., Johnson, M.H. & Karmiloff-Smith, A., (2000). 'Electrophysiological correlates of cross-linguistic speech perception in native English speakers'. *Behavioral Brain Research* 111 (1-2):13-23.
- Rosburg, T., Trautner, P., Dietl, T., Korzyukov, O., Boutros, N., Schaller, C., Elger, C. & Kurthen, M., (2005). 'Subdural recordings of the mismatch negativity (MMN) in patients with focal epilepsy'. *Brain*, 128 (4):819-828.
- Rugg, M. & Coles, M., (1996). *Electrophysiology of Mind: Event-related Brain Potentials and Cognition*. Oxford University Press, USA.
- Saarinen, J., Paavilainen, P., Schröger, E., Tervaniemi, M. & Näätänen, R., (1992). 'Representation of abstract stimulus attributes in human brain'. *NeuroReport*, 3:1149-1151.
- Sable, J., Gratton, G. & Fabiani, M., (2003). 'Sound presentation rate is represented logarithmically in human cortex'. *European Journal of Neuroscience*, 17 (11):2492-2496.
- Sagey, E., (1986). 'The representation of features and relations in non-linear phonology' [*Dissertation*]. Massachusetts Institute of Technology.
- Sams, M., Aulanko, R., Aaltonen, O. & Näätänen, R., (1990). 'Event-related potentials to infrequent changes in synthesized phonetic stimuli'. *Journal of Cognitive Neuroscience*, 2 (4):344-357.

- Sams, M., Paavilainen, P., Alho, K. & Näätänen, R., (1985a). 'Auditory frequency discrimination and event-related potentials'. *Electroencephalography and Clinical Neurophysiology*, 62 (6):437-448.
- Salthouse, T., (1996). 'The processing-speed theory of adult age differences in cognition'. *Psychological Review*, 103 (3):403-427.
- Salthouse, T., (2000). 'Aging and measures of processing speed'. *Biological Psychology*, 54 (1-3):35-54.
- Scharinger, M., (2009). 'Minimal representations of alternating vowels'. *Lingua*, 119 (10):1414-1425.
- Scharinger, M., & Lahiri, A. (2010). Height differences in English dialects: Consequences for processing and representation. *Language and Speech*, 53(2):245-272.
- Scharinger, M., Reetz, H. & Lahiri, A., (2009). 'Levels of regularity in inflected word form processing'. *The Mental Lexicon*, 4 (1):77-114.
- Scherg, M., Vajsar, J. & Picton, T., (1989). 'A source analysis of the late human auditory evoked potentials'. *Journal of Cognitive Neuroscience*, 1 (4):336-355.
- Schröger, E., (1996). 'A neural mechanism for involuntary attention shifts to changes in auditory stimulation'. *Journal of Cognitive Neuroscience*, 8 (6):527-539.
- Schröger, E., (1997). 'On the detection of auditory deviations: A pre-attentive activation model'. *Psychophysiology*, 34 (3):245-257.
- Schröger, E., (1998). 'Measurement and interpretation of the mismatch negativity'. *Behavior Research Methods, Instruments & Computers*, 30: 131-145.
- Schröger, E., Giard, M.H., & Wolff, C., (2000). Auditory distraction: event-related potential and behavioural indices. *Clinical Neurophysiology*, 111, 1450–1460.
- Schröger, E., Paavilainen, P. & Näätänen, R., (1994). 'Mismatch negativity to changes in a continuous tone with regularly varying frequencies'. *Electroencephalography and Clinical Neurophysiology*, 92:140-140.
- Schröger, E. & Wolff, C., (1996). 'Mismatch response of the human brain to changes in sound location'. *NeuroReport*, 7 (18):3005-3008.
- Schröger, E., & Wolff, C., (1998a). Behavioural and electrophysiological effects of taskirrelevant sound change: a new distraction paradigm. *Cognitive Brain Research*, 7, 71–87.
- Schröger, E., & Wolff, C., (1998b). Attentional orienting and reorienting is indicated by human event-related brain potentials. *Neuroreport*, 9, 3355–3358.
- Shafer, V.L., Schwartz, R.G. & Kurtzberg, D., (2004). 'Language-specific memory traces of consonants in the brain'. *Brain Research. Cognitive Brain Research*, 18 (3):242-254.
- Sharma, A. & Dorman, M., (1999). 'Cortical auditory evoked potential correlates of categorical perception of voice-onset time'. *The Journal of the Acoustical Society of America*, 106:1078.

- Sharma, A., Kraus, N., McGee, T., Carrell, T. & Nicol, T., (1993). 'Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential'. *Electroencephalography and Clinical Neurophysiology*, 88 (1):64-71.
- Shelley, A., Ward, P., Catts, S., Michie, P., Andrews, S. & McConaghy, N., (1991). 'Mismatch negativity: an index of a preattentive processing deficit in schizophrenia'. *Biological Psychiatry*, 30 (10):1059-1062.
- Shestakova, A., Brattico, E., Huotilainen, M., Galunov, V., Soloviev, A., Sams, M., Ilmoniemi, R.J. & Näätänen, R., (2002). 'Abstract phoneme representations in the left temporal cortex: Magnetic mismatch negativity study'. *Neuroreport: For Rapid Communication of Neuroscience Research*, 13 (14):1813-1816.
- Shtyrov, Y., Hauk, O. & Pulvermüller, F., (2004). 'Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words'. *European Journal of Neuroscience*, 19 (4):1083-1092.
- Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R.J. & Näätänen, R., (2000). 'Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres.'. *NeuroImage*, 12 (6): 657-663.
- Shtyrov, Y. & Pulvermüller, F., (2002a). 'Memory traces for inflectional affixes as shown by mismatch negativity'. *European Journal of Neuroscience*, 15 (6):1085-1091.
- Shtyrov, Y. & Pulvermüller, F., (2002b). 'Neurophysiological evidence of memory traces for words in the human brain'. *Cognitive Neuroscience and Neuropsychology*, 13 (4):521-526.
- Shtyrov, Y. & Pulvermüller, F., (2007a). 'Early MEG activation dynamics in the left temporal and inferior frontal cortex reflect semantic context integration'. *Journal of Cognitive Neuroscience*, 19 (10):1633-1642.
- Shtyrov, Y. & Pulvermüller, F., (2007b). 'Language in the mismatch negativity design: Motivations, benefits, and prospects'. *Journal of Psychophysiology*, 21 (3/4):176-187.
- Shtyrov, Y., Pulvermüller, F., Näätänen, R. & Ilmoniemi, R., (2003). 'Grammar processing outside the focus of attention: An MEG study'. *Journal of Cognitive Neuroscience*, 15 (8):1195-1206.
- Shtyrov, Y., Osswald, K. & Pulvermüller, F., (2008). 'Memory traces for spoken words in the brain as revealed by the hemodynamic correlate of the mismatch negativity'. *Cerebral Cortex*, 18 (1):29-37.
- Simons, R., Graham, F., Miles, M. & Chen, X., (2001). 'On the relationship of P3a and the Novelty-P3'. *Biological Psychology*, 56 (3):207-218.
- Sokolov, E., (1963). *Perception and the Conditioned Reflex*. New York: Pergamon Press.
- Soltani, M. & Knight, R., (2000). 'Neural Origins of the P300'. *Critical Reviews in Neurobiology*, 14 (3-4):199-224.
- Spencer, K., Dien, J. & Donchin, E., (1999). 'A componential analysis of the ERP elicited by novel events using a dense electrode array'. *Psychophysiology*, 36 (03):409-414.

- Spencer, K., Dien, J. & Donchin, E., (2001). 'Spatiotemporal analysis of the late ERP responses to deviant stimuli'. *Psychophysiology*, 38 (02):343-358.
- Squires, N., Squires, K. & Hillyard, S., (1975). 'Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man'. *Electroencephalography and Clinical Neurophysiology*, 38 (4):387-401.
- Stevens, K.N., (1986). 'Models of phonetic recognition II: A feature-based model of speech recognition'. In: Mermelstein, P. (Ed). *Proceedings of the Montreal Saeellite Symposium on Speech Recognition*, Twelfth International Congress on Acoustics.
- Stevens, K., (2002). 'Toward a model for lexical access based on acoustic landmarks and distinctive features'. *The Journal of the Acoustical Society of America*, 111:1872.
- Stevens, K. & Keyser, S., (2008). 'Quantal theory, enhancement and overlap'. *Journal of Phonetics*.
- Sussman, E., Ritter, W. & Vaughan, H., (1999). 'An investigation of the auditory streaming effect using event-related brain potentials'. *Psychophysiology*, 36 (01):22-34.
- Sutton, S., Braren, M., Zubin, J. & John, E., (1965). 'Evoked-potential correlates of stimulus uncertainty'. *Science*, 150 (3700):1187-1188.
- Tervaniemi, M., Medvedev, S., Alho, K., Pakhomov, S., Roudas, M., Van Zuijen, T. & Näätänen, R., (2000). 'Lateralized automatic auditory processing of phonetic versus musical information: a PET study'. *Human Brain Mapping*, 10 (2):74-79.
- Tiitinen, H., May, P., Reinikainen, K. & Näätänen, R., (1994). 'Attentive novelty detection in humans is governed by pre-attentive sensory memory'. *Nature*, 372.
- Titova, N. & Näätänen, R., (2001). 'Preattentive voice discrimination by the human brain as indexed by the mismatch negativity'. *Neuroscience Letters*, 308 (1):63-65.
- Toiviainen, P., Tervaniemi, M., Louhivuori, J., Sailer, M., Huottilainen, M. & Näätänen, R., (1998). 'Timbre similarity: Convergence of neural, behavioral, and computational approaches'. *Music Perception*, 16:223-242.
- Tsuchiya, H., Yamaguchi, S., & Kobayashi, S., (2000). Impaired novelty detection and frontal lobe dysfunction in Parkinson's disease. *Neuropsychologia*, 38, 645-654.
- Uwer, R., Albrecht, R. & Von Suchodoletz, W., (2002). 'Automatic processing of tones and speech stimuli in children with specific language impairment'. *Developmental Medicine and Child Neurology*, 44 (08):527-532.
- Verhaeghen, P. & De Meersman, L., (1998). 'Aging and the Stroop effect: A meta-analysis'. *Psychology and Aging*, 13:120-126.
- Verleger, R., 1988. A critique of the context updating hypothesis and an alternative interpretation of P3. *Behavioral and Brain Sciences*, 11, 343-427.
- Verleger, R., & Hopmann, G. (1989). When things get exciting: Is P3 modulated by background suspense-related negativity? *Journal of Psychophysiology*, 3, 269-279.
- Verleger, R., Jaskowski, P. & Wascher, E., (2005). 'Evidence for an integrative role of P3b in linking reaction to perception'. *Journal of Psychophysiology*, 19 (3):165-181.

- Walter, M. & Hacquard, V., (2004). 'MEG evidence for phonological underspecification'. MIT Phonology Circle.
- Walter, W., Cooper, R., Aldridge, V., McCallum, W. & Winter, A., (1964). 'Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain'. *Nature*, 203:380-384.
- Weber, C., Hahne, A., Friedrich, M. & Friederici, A., (2004). 'Discrimination of word stress in early infant perception: Electrophysiological evidence'. *Cognitive Brain Research*, 18 (2):149-161.
- Wheeldon, L. & Waksler, R., (2004). 'Phonological underspecification and mapping mechanisms in the speech recognition lexicon'. *Brain and Language*, 90 (1-3):401-412.
- Wiese, R., (2000). *The Phonology of German*. Oxford: Oxford University Press.
- Winkler, I. & Czigler, I., (1998). 'Mismatch negativity: deviance detection or the maintenance of the 'standard''. *NeuroReport*, 9 (17):3809-3813.
- Winkler, I., Cowan, N., Csépe, V., Czigler, I. & Näätänen, R., (1996). 'Interactions between transient and long-term auditory memory as reflected by the mismatch negativity'. *Journal of Cognitive Neuroscience*, 8 (5):403-415.
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R.J. & Näätänen, R., (1999a). 'Brain responses reveal the learning of foreign language phonemes'. *Psychophysiology*, 36 (5):638-642.
- Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csepe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A. & Näätänen, R., (1999b). 'Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations'. *Cognitive Brain Research*, 7 (3):357-369.
- Winkler I. & Näätänen R., (1993). Event-related brain potentials to infrequent partial omissions in series of auditory stimuli. In: H.-J. Heinze, T.F. Münte, & G.R. Mangun, (Eds.) *New Developments in Event-Related Potentials* (pp. 219-26). Boston-Basel-Berlin: Birkhäuser.
- Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M. & Näätänen, R., (1990). 'The effect of small variation of the frequent auditory stimulus on the event-related brain potential to the infrequent stimulus'. *Psychophysiology*, 27 (2):228-235.
- Winkler, I., Reinikainen, K. & Näätänen, R., (1993). 'Event-related brain potentials reflect traces of echoic memory in humans'. *Perception and Psychophysics*, 53 (4):443-449.
- Wunderlich, J.L. & Cone-Wesson, B.K., (2001). 'Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential'. *Journal of the Acoustical Society of America*, 109 (4):1526-1537.
- Yasin, I., (2007). 'Hemispheric differences in processing dichotic meaningful and non-meaningful words'. *Neuropsychologia*, 45 (12):2718-2729.
- Yago, E., Corral, M. & Escera, C., (2001a). 'Activation of Brain Mechanisms of involuntary

Attention to different Magnitudes of auditory Frequency Deviation'. <http://www.uni-leipzig.de/~psycho/fechner/generalinfo/PDFs/EYago.pdf>.

Zimmerer, F., Reetz, H. & Lahiri, A., (2009). 'Place assimilation across words in running speech: Corpus analysis and perception'. *The Journal of the Acoustical Society of America*, 125:2307.

Zimmerer, F., (2009). 'Reduction in natural speech' [*Dissertation*]. Frankfurt: Goethe University of Frankfurt.