รายงานวิจัยฉบับสมบูรณ์

โครงการวิจัยการศึกษาการทำงานของสมองมนุษย์ด้านการรับฟังเสียงในการรับรู้ระบบคำในภาษาไทย

โดย ดร.วิเชียร สิทธิประภาพร และคณะ

1 กรกฎาคม 2548

รายงานวิจัยฉบับสมบูรณ์

โครงการ การศึกษาการทำงานของสมองมนุษย์ ด้านการรับฟังเสียงในการรับรู้ระบบคำในภาษาไทย

คณะผู้วิจัย	สังกัด
ดร.วิเชียร สิทธิประภาพร	มหาวิทยาลัยมหิดล
ผศ.น.พ.จิตตินทร์ จินดาดวงรัตน์	มหาวิทยาลัยมหิดล
รศ.ดร.นัยพินิจ คชภักดี	มหาวิทยาลัยมหิดล

สนับสนุนโดย สำนักงานคณะกรรมการการอุดมศึกษา และสำนักงานกองทุนสนับสนุนการวิจัย

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude and deep appreciation to my mentors, Associate Professor Dr. Naiphinich Kotchabhakdi and Dr. Mari Tervaniemi, for their insightful advice, encouragement and support throughout the course of this investigation. In addition, I am equally special grateful to Wilairat Khampan, Tsui-Chin Wang, Pilant Ananchaipatthana, for their assistance during the volunteer arrangements. It is a great pleasure to acknowledge the help of many individuals, without whose help this research could not have been written. First and foremost, I am indebted to Dr. Mari Tervaniemi, for her invaluable assistance during this research effort.

I would like to thank most sincerely the following people, without whom this research would not be what it is: Professor Dr. Wichai Boonsaeng and Professor Dr. Kanok Phawwanawisutthiphaisarn. Finally, I would like to give tribute to my family who forgave me my absences, tolerated my depressed times, and shared with me my exhilaration as this research moved through its various phases.

The studies of this research were financially supported by the Thailand Research Fund (TRF) under grant number MRG4780099, Thailand. All experimental works were carried out in cooperation with the Neuro-Behavioural Biology Center of Institute of Science and Technology for Research and Development, Mahidol University, Salaya, Nakhonpathom, Thailand.

Wichian Sittiprapaporn, Ph.D.

บทคัดย่อ

ในการศึกษาด้านการทำงานของสมองในการจำแนกความแตกต่างของหน่วยเสียงในภาษานั้น คลื่นของสมองที่เรียกว่า มีสแมทเนกาทิวิตี้ (Mismatch Negativity, MMN) ซึ่งเป็นคลื่นของสมองที่ได้จากการจำแนกเสียงต่าง ๆทั้งที่เป็นเสียงในภาษาและไม่เป็นเสียงในภาษาในขณะที่ไม่ตั้งใจ (Unattended Auditory Perception) ซึ่งมีสแมทเนกาทิวิตี้ดังกล่าว ได้ถูกนำมาใช้กันอย่างแพร่หลายโดยเฉพาะอย่างยิ่งนักประสาทวิทยาศาสตร์ที่มุ่งศึกษาในด้านการทำงานของสมองที่เกี่ยวข้องกับภาษา (Neuro-linguistics) โดยส่วนใหญ่จะมุ่งเน้นศึกษาเฉพาะกลุ่มของภาษาที่ไม่ใช้เสียงวรรณยุกต์ในการกำหนดความหมายของคำ อาทิเช่น ภาษาอังกฤษเป็นต้น ซึ่งปัจจุบัน ยังไม่ปรากฏว่ามีการนำเอา มีสแมทเนกาทิวิตี้ มาใช้ในการศึกษาถึงการทำงานของสมองของกลุ่มผู้ที่ใช้ภาษาที่ใช้วรรณยุกย์ในการสื่อสารแต่อย่างใด ดังนั้น การศึกษาวิจัยฉบับนี้จึงได้ทำการศึกษาว่า จะสามารถนำเอา มีสแมชเนกาทิวิตี้ดังกล่าว มาใช้ในการศึกษาการทำงานสมองของอาสาสมัครทั้งชาวไทยและชาวต่างชาติ ในการจำแนกระบบคำในภาษาไทยได้อย่างไร ในการศึกษาครั้งนี้ ผู้วิจัยได้ใช้หน่วยเสียงต่าง ๆที่มีความแตกต่างกันทั้งทางด้านพยัญชนะ สระ และวรรณยุกต์ ที่ประกอบกันภายในโครงสร้างของระบบคำในภาษาไทย ซึ่งสามารถแบ่งออกได้เป็น 3 กลุ่ม ตามวัตถุประสงค์ของการศึกษา ดังนี้ (1) การศึกษาที่เกี่ยวกับการจำแนกประเภทของสระ (Category of Vowel) ประกอบด้วย (ก) คำว่า /ปี/ มีเสียงสระส่วนหน้า ไม่ห่อปากและวรรณยุกต์เสียงกลางระดับ (ข) คำว่า /โป/ มีเสียงสระส่วนหน้า ห่อปากและวรรณยุกต์เสียงกลางระดับ และ (ค) คำว่า /ปอ/ มีเสียงสระส่วนหลัง ห่อปากและวรรณยุกต์เสียงกลางระดับ (2) การศึกษาที่เกี่ยวกับการจำแนกความแตกต่างทางด้านเสียงวรรณยุกต์ (Tone Change) และความยาวของสระ (Vowel Duration) ประกอบด้วย (ก) คำว่า /คาม/ มีวรรณ์ยุกต์เสียงกลางระดับ (เสียงสามัญ) และสระเสียงยาว (ข) คำว่า /ข้าม/ มีวรรณยุกต์เสียงสูงตก (เสียงโท) และสระเสียงยาว (ค) คำว่า /ขาม/ มีวรรณยุกต์เสียงต่ำขึ้น (เสียงจัตวา) และสระเสียงยาว (ง) คำว่า /คำ/ มีวรรณยุกต์เสียงกลางระดับ (เสียงสามัญ) และสระเสียงสั้น (จ) คำว่า /ค่ำ/ มี

วรรณยุกต์เสียงสูงตก (เสียงโท) และสระเสียงสั้น (ฉ) คำว่า /ขำ/ มีวรรณยุกต์เสียงต่ำขึ้น (เสียงจัตวา) และสระเสียงสั้น และ (3) การศึกษาที่เกี่ยวกับการจำแนกประเภทของพยัญชนะชนิดควบกล้ำ (Cluster Consonant) และไม่ควบกล้ำ (Noncluster Cluster Consonant) ประกอบด้วย (ก) คำว่า /กาง/ มีเสียงพยัญชนะชนิดไม่ควบกล้ำและมีวรรณยุกต์เสียงกลางระดับ (เสียงสามัญ) และ (ข) คำว่า /กลาง/ มีเสียงพยัญชนะชนิดควบกล้ำและมีวรรณยุกต์เสียงกลางระดับ (เสียงสามัญ)

ในการศึกษาครั้งนี้ ผู้วิจัยได้ทำการตรวจวัดการทำงานของสมองในรูปแบบของคลื่นสมองในขณะที่สมองกำลังจำแนกความแตกต่างของหน่วยเสียงดังกล่าว โดยคลื่นสมองที่ได้นั้นจะมาจากการวัดสัญญาณจากอิเล็คโทดจำนวน 21 อิเล็คโทรด ที่ได้ติดตามบริเวณต่าง ๆของศีรษะของอาสาสมัคร จากการศึกษาพบว่าสมองมีการทำงานที่เด่นทั้งสองซีกในด้านต่าง ๆดังต่อไปนี้

(1) การจำแนกสระที่ต่างประเภทกัน เช่น สระส่วนหน้า ไม่ห่อปาก และสระส่วนหลัง ห่อปาก การจำแนกสระเสียงสั้นออกจากสระเสียงยาว

(2) การจำแนกการเปลี่ยนแปลงของสระจากสระเสียงยาวไปเป็นสระเสียงสั้นได้ดีกว่าการเปลี่ยนจากสระเสียงสั้นไปเสียงยาว

(3) การจำแนกการเปลี่ยนแปลงของเสียงวรรณยุกย์จากเสียงระดับไปเป็นเสียงต่างระดับ เช่น จากเสียงวรรณยุกต์กลางระดับ (เสียงสามัญ) ไปเป็นเสียงสูงตก (เสียงโท) แงะเสียงต่ำขึ้น (เสียงจัตวา)

(4) การจำแนกพยัญชนะเสียงควบกล้ำออกจากพยัญชนะเสียงไม่ควบกล้ำ

ทั้งนี้ ความสามารถในการทำงานของสมองทั้งสองซีกดังกล่าวสามารถพบได้ในอาสาสมัครทั้งสองกลุ่ม โดยเฉพาะอย่างยิ่ง คลื่นของสมองของอาสาสมัครชาวไทยจะเด่นกว่าคลื่นของสมองของชาวต่างชาติ ดังนั้น จากผลการศึกษาทำให้ทราบว่า มีสแมทเนกาทิวิตี้ สามารถนำมา

ใช้ในการตรวจวัดการทำงานของสมองในระบบคำในภาษาไทยด้านต่าง ๆดังต่อไปนี้ได้ คือ (ก) การจำแนกประเภทของสระ (ข) การจำแนกความแตกต่างทางด้านความสั้นยาวของสระ (ค) การจำแนกความแตกต่างทางด้านวรรณยุกต์ และ (ง) การจำแนกประเภทของพยัญชนะชนิดควบกล้ำและไม่ควบกล้ำ เป็นต้น ทั้งนี้ ความสามารถของสมองดังกล่าวจะใช้เวลาในการจำแนกเสียงตั้งแต่ระยะเวลา 2.5 ถึง 4.2 วินาที (150-250 msec) ตามแต่ละประเภทของคำ ดังนั้น จึงอาจกล่าวได้ว่า ค่ามีสแมทเนกาทิวิตี้ ที่ได้จากการจำแนกระบบคำในภาษาไทยสามารถนำมาใช้เป็นเครื่องชี้วัดความสามารถของสมองของอาสาสมัครในการเรียนรู้ภาษาได้ทั้งในระดับที่ตั้งใจและไม่ตั้งใจ (Attended and Unattended Auditory Processing) สมองทั้งสองซีก ทั้งนี้กลไกการทำงานแบบอัตมัติดังกล่าว จึงสามารถนำไปใช้เป็นหลักเกณฑ์สากลในการจำแนกเสียงในภาษาของทั้งที่ใช้วรรณยุกต์และไม่ใช้วรรณยุกต์ในการกำหนดความหมายของคำ

ABSTRACT

Previous event-related brain potential research showed that mismatch negativity (MMN) was elicited by phoneme contrasts. The present study tested whether the MMN would be elicited in response to temporal and spectral changes in Thai synthesized words for both native- and non-speakers of Thai. Stimuli included three computer-synthesized speech with varying duration or frequency modulation depending on each main experiment: Category of vowel perception: (1) /pi/ - front unrounded vowel and level tone, (2) /po/ - front rounded vowel and level tone, (3) /pO/ - back rounded vowel and level tone; Vowel duration and tone changes perception: (4) /k^{h}aam/ - long vowel duration and level tone, (5) /k^{h}aam/ - long vowel duration and falling tone, (6) /k^{h}aam/ - long vowel duration and rising tone, (7) /k^{h}am/ - short vowel duration and level tone, (8) /k^{h}am/ - short vowel duration and falling tone, (9) /k^{h}am/ - short vowel duration and rising tone; Cluster consonant perception: (10) /kaang/ - noncluster initial consonant and level tone, (11) /klaang/ - cluster initial consonant and level tone. Electroencephalography responses were recorded with a 21 active electrodes. Our recent MMN studies found that across-category of vowel, long-to-short duration; level-to-falling/rising tone and noncluster-to-cluster consonant changes elicited a prominent MMN bilaterally for both groups, unlike within-category of vowel, short-to-long vowel duration, falling/rising-to-level tone, and cluster-to-noncluster consonant changes. In other word, the shortened-vowel duration changes and tone changes in Thai words elicited a prominent MMN in two hemispheres for both groups. The results showed that the prominent MMN component was generated in the following aspects: (1) within-to-across category of vowel, (2) long-to-short duration changes, level-to-falling/rising changes and rising-to-falling tone changes, (3) noncluster-to-cluster consonant changes (peaking around 150-250 ms) in each hemisphere for both native and nonnative subjects. The MMN component in Thai words was particularly more sensitive to across-category of vowel, duration shortening and contour tone changes, and cluster consonant. In summary, changes in duration shortening contour tone, and cluster consonant changes are particularly salient cues for pre-attentive auditory change detection in each hemisphere. Automatic detection of changes in the above physiological detection is a useful index of language-non-specific auditory memory traces.

หน้าสรุปโครงการ (Executive Summary)

ทุนพัฒนาศักยภาพในการทำงานวิจัยของอาจารย์รุ่นใหม่

1. ชื่อโครงการ (ภาษาไทย) การศึกษาการทำงานของสมองมนุษย์ด้านการรับฟังเสียงในการรับรู้ระบบคำในภาษาไทย

(ภาษาอังกฤษ) Functional Specialization of the Human Auditory Cortex for Processing of Lexicality in Thai Language

2. ชื่อหัวหน้าโครงการวิจัย

ดร.วิเชียร สิทธิประภาพร
โครงการวิจัยชีววิทยาระบบประสาทและพฤติกรรม
สถาบันวิจัยและพัฒนาวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยมหิดล ศาลายา นครปฐม 73170
โทรศัพท์: 0-2441-9321 โทรสาร: 0-2441-9743
e-mail: stwst@mahidol.ac.th

3. สาขาวิชาที่ทำการวิจัย: สาขาวิทยาศาสตร์การแพทย์

4. งบประมาณทั้งโครงการ: 240,000 บาท

5. ระยะเวลาดำเนินงาน: 1 ปี

6. แหล่งทุนสนับสนุนการวิจัย: สำนักงานกองทุนสนับสนุนการวิจัย (สกว.)

รหัส MRG4780099

7. ปัญหาที่ทำการวิจัย และความสำคัญของปัญหา

โครงการวิจัยศึกษาความสัมพันธ์ระหว่างการทำงานของสมองกับการรับรู้คำโครงสร้างพยางค์เดียว (Monosyllabic word) ได้จัดทำขึ้นโดยมีวัตถุประสงค์เพื่อศึกษาบริเวณของสมองและความสัมพันธ์ของการทำงานสมองบริเวณดังกล่าวในประเภทและความสั้นยาวของหน่วยเสียงสระ การเปลี่ยนแปลงของหน่วยเสียงวรรณยุกต์ ตลอดจนหน่วยเสียงพยัญชนะควบกล้ำและไม่ควบกล้ำในระบบคำที่เป็นโครงสร้างพยางค์เดียวในภาษาไทย โดยการศึกษาลักษณะของคลื่นสมองที่เกิดขึ้นจากการทำงานของสมองในขณะแยกหน่วยเสียงดังกล่าว อีกทั้งยังศึกษาถึงผลของเสียงพูดต่อการทำงานของสมองที่แสดงออกในรูปของคลื่นสมอง ข้อมูลที่อยู่ในรูปของคลื่นสมองได้นำไปวิเคราะห์และแสดงผลในรูปของภาพการทำงานของสมอง (Neuroimaging) บริเวณสมองส่วนที่ทำหน้าที่ในการแยกหน่วยเสียงดังกล่าวในคำโครงสร้างพยางค์เดียว

ข้อมูลที่ได้จากการศึกษาวิจัยครั้งนี้สามารถนำไปใช้เป็นประโยชน์ในการศึกษาและเข้าใจถึงกลไกของสมองในการรับรู้หน่วยเสียงดังกล่าวในคำโครงสร้างพยางค์เดียว ของภาษาที่อยู่ในกลุ่มที่ใช้วรรณยุกต์ (Tonal Languages) ในการกำหนดความหมายของคำ คือ ภาษาไทย อีกทั้งยังทำให้เข้าใจถึงกลไกของสมองในการแยกแยะเสียงหรือคำพูดที่มาจากเสียงของเจ้าของภาษาเอง (Native Language) ทั้งนี้ก็เพื่อที่จะเป็นแนวทางในการเรียนรู้ภาษาได้อย่างถูกต้อง ตลอดจนเพื่อนำผลการศึกษาการทำงานของบริเวณสมองในการแยกแยะหน่วยเสียงดังกล่าวในคำโครงสร้างพยางค์เดียวที่ได้จากอาสาสมัคร หรือกลุ่มที่มีสุขภาพดี สมบูรณ์แข็งแรง ไปเป็นข้อมูลพื้นฐานในการเปรียบเทียบกับกลุ่มผู้ป่วยที่มีความผิดปรกติ หรือผู้ป่วยที่มีความบกพร่องในการแยกแยะหน่วยเสียงดังกล่าวและการเข้าใจความหมายของคำในภาษา อีกทั้งผลของการศึกษายังได้นำไปประยุกต์เป็นแนวทางในการรักษากลุ่มผู้ป่วยที่มีความบกพร่องในการแยกแยะเสียงและความหมายของคำในภาษาดังกล่าว เพื่อให้ผู้ป่วยสามารถสื่อสารและดำรงชีวิตอยู่ในสังคมกับบุคคลปรกติทั่วไปได้ ตลอดจนการนำผลการศึกษาไปเป็นแนวทางในการส่งเสริม

พัฒนาการและการรักษาภาวะการเจริญเติบโตล่าช้าของเด็กที่มีความบกพร่องทางพัฒนาการในด้านการแยกแยะเสียงและความหมายของคำ รวมไปถึงพัฒนาการในด้านทักษะฟัง พูด อ่าน และเขียน ทั้งในภาษาตนเอง และภาษาที่สอง

8. วัตถุประสงค์

8.1. เพื่อศึกษากลไกการทำงานของสมองและความสัมพันธ์ของสมองกับการแยกแยะประเภทและความสั้นยาวของหน่วยเสียงสระ การเปลี่ยนแปลงของหน่วยเสียงวรรณยุกต์ ตลอดจนหน่วยเสียงพยัญชนะควบกล้ำและไม่ควบกล้ำในระบบคำที่เป็นโครงสร้างพยางค์เดียวในภาษาไทย

8.2. เพื่อศึกษาบริเวณของสมองส่วนที่เกี่ยวข้องกับการตอบรับต่อการแยกแยะประเภทและความสั้นยาวของหน่วยเสียงสระ การเปลี่ยนแปลงของหน่วยเสียงวรรณยุกต์ ตลอดจนหน่วยเสียงพยัญชนะควบกล้ำและไม่ควบกล้ำไปทำหน้าที่ในการกำหนดความหมายของคำในภาษาไทย

8.3. เพื่อศึกษาลักษณะของคลื่นสมองที่เกิดขึ้นจากการทำงานของสมองในขณะแยกแยะประเภทและความสั้นยาวของหน่วยเสียงสระ การเปลี่ยนแปลงของหน่วยเสียงวรรณยุกต์ ตลอดจนหน่วยเสียงพยัญชนะควบกล้ำและไม่ควบกล้ำในระบบคำที่เป็นโครงสร้างพยางค์เดียวในภาษาไทย

8.4. เพื่อศึกษาถึงผลของเสียงพูดที่เป็นเสียงของเจ้าของภาษา ต่อการทำงานของสมองที่แสดงออกในรูปของคลื่นสมอง

8.5. เพื่อนำผลการศึกษาที่ได้ไปประยุกต์เป็นแนวทางในการรักษากลุ่มผู้ป่วยที่มีความบกพร่องในการแยกแยะเสียงและความหมายของคำในภาษา เพื่อให้สามารถสื่อสาร และ

ดำรงชีวิตอยู่ในสังคมกับบุคคลปรกติทั่วไปได้ ตลอดจนเพื่อนำผลการศึกษาไปเป็นแนวทางในการส่งเสริมพัฒนาการและการรักษาภาวะการเจริญเติบโตล่าช้าของเด็กที่มีความบกพร่องทางพัฒนาการในด้านการแยกแยะเสียงและความหมายของคำ รวมไปถึงพัฒนาการในด้านทักษะฟัง พูด อ่าน และเขียน ทั้งในภาษาตนเอง และภาษาที่สอง

9. ประโยชน์ที่ได้รับจากการวิจัย

9.1. ทำให้เข้าใจกลไกของสมองมนุษย์ในด้านการรับฟังเสียงและการรับรู้ประเภทและความสั้นยาวของหน่วยเสียงสระ การเปลี่ยนแปลงของหน่วยเสียงวรรณยุกต์ ตลอดจนหน่วยเสียงพยัญชนะควบกล้ำและไม่ควบกล้ำในระบบคำที่เป็นโครงสร้างพยางค์เดียวในภาษาที่อยู่ในกลุ่มที่ใช้วรรณยุกต์ (Tonal Languages) ในการกำหนดความหมายของคำ คือ ภาษาไทย

9.2. ทำให้เข้าใจกลไกการทำงานของสมองและบริเวณของสมองที่เกี่ยวข้องในการการแยกแยะประเภทและความสั้นยาวของหน่วยเสียงสระ การเปลี่ยนแปลงของหน่วยเสียงวรรณยุกต์ ตลอดจนหน่วยเสียงพยัญชนะควบกล้ำและไม่ควบกล้ำในระบบคำที่เป็นโครงสร้างพยางค์เดียวในภาษาที่อยู่ในกลุ่มที่ใช้วรรณยุกต์ (Tonal Languages) ในการกำหนดความหมายของคำ คือ ภาษาไทย และเป็นเสียงพูดที่มาจากเสียงของเจ้าของภาษาเอง

9.3. สามารถนำผลการศึกษากลไกการทำงานของสมองในการแยกแยะประเภทและความสั้นยาวของหน่วยเสียงสระ การเปลี่ยนแปลงของหน่วยเสียงวรรณยุกต์ ตลอดจนหน่วยเสียงพยัญชนะควบกล้ำและไม่ควบกล้ำในระบบคำที่เป็นโครงสร้างพยางค์เดียวในภาษาไทยที่ได้จากอาสาสมัครหรือกลุ่มที่มีสุขภาพดี สมบูรณ์แข็งแรง ไปเป็นแนวทางพื้นฐาน (Basic Control) ในการเปรียบเทียบกับกลุ่มผู้ป่วยที่มีความผิดปรกติ หรือผู้ป่วยที่มีความบกพร่องในการแยกแยะเสียงหรือคำพูดในด้านการรับฟังเสียงและการรับรู้ระบบคำที่ประกอบด้วยหน่วยเสียงดังกล่าว ในโครงการวิจัยต่อไป

9.4. สามารถนำผลการศึกษาที่ได้ไปประยุกต์เป็นแนวทางในการรักษากลุ่มผู้ป่วยที่มีความบกพร่องในการแยกแยะเสียงและความหมายของคำในภาษา เพื่อให้สามารถสื่อสารและดำรงชีวิตอยู่ในสังคมกับบุคคลปรกติทั่วไปได้

9.5. สามารถนำผลการศึกษาไปเป็นแนวทางในการส่งเสริมพัฒนาการและการรักษาภาวะการเจริญเติบโตล่าช้าของเด็กที่มีความบกพร่องทางพัฒนาการในการแยกแยะเสียงและความหมายของคำในภาษา ตลอดจนพัฒนาการในด้านทักษะฟัง พูด อ่าน และเขียน

9.6. ได้นำผลการวิจัยไปนำเสนอผลงานในรูปของการประชุมวิชาการด้านประสาทวิทยาศาสตร์และในรูปของการตีพิมพ์ผลงานในวารสารระดับนานาชาติที่มีผู้ประเมิน (Peer Reviews) และมีค่า impact factor ดังต่อไปนี้

Contributions to Peer-Reviews Journals

ปีที่ 2 ลำดับที่ 1

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C., and Kotchabhakdi, N. Preattentive Discrimination of Across-category and Within-category Change in Consonant-Vowel Syllable. *NeuroReport* 2005; 16(13): 1513-1518. [ค่า Impact Factor = 2.503]

ปีที่ 2 ลำดับที่ 2

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C. and Kotchabhakdi, N. Electrophysiological Measurement of Vowel Duration Changes in Monosyllabic Thai Words. *NeuroReport* 2005 (Submitted). [ค่า Impact Factor = 2.503]

ปีที่ 2 ลำดับที่ 3

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C. and Kotchabhakdi, N. Auditory Preattentive Processing of Cluster vs. Noncluster Initial Consonant Perception in Monosyllabic Thai Words: Preliminary Study. *NeuroReport* 2005 (Submitted). [ค่า Impact Factor = 2.503]

ปีที่ 2 ลำดับที่ 4

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C. and Kotchabhakdi, N. Functional Specialization of the Human Auditory Cortex for Processing of Lexicality in Thai Language *Cognitive Brain Research* (In Preparation).

Published Abstracts

ปีที่ 1 ลำดับที่ 1

Awikunparsert, C., Sittiprapaporn, W., Chindaduangratn, C., and Kotchabhakdi, N. Neurophysiological Evidence of Memory Traces for Words with Vowel of Lexical Tones. *International Journal of Behavioral Medicine* 2004; 11:107. [ค่า Impact Factor = 1.39]

ปีที่ 1 ลำดับที่ 2

Sittiprapaporn, W., Awikunparsert, C., Chindaduangratn, C., and Kotchabhakdi, N. Auditory Processing for Mismatch Negativity Responses to Across-and Within-category Change of Vowels. *International Journal of Behavioral Medicine* 2004; 11:111. [ค่า Impact Factor = 1.39]

ปีที่ 1 ลำดับที่ 3

Sittiprapaporn, W., Awikunparsert, C., Chindaduangratn, C., and Kotchabhakdi, N. Memory Traces for Cluster vs. Noncluster Consonants as Shown by Mismatch Negativity. *International Journal of Behavioral Medicine* 2004; 11:111. [ค่า Impact Factor = 1.39]

ปีที่ 2 ลำดับที่ 1

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. Source Generator of Across-and Within-Category Changes of Vowel in Consonant-Vowel (CV) Syllable: Low-resolution Electromagnetic Tomography (LORETA) Study. *Brain Topography* 2005 (in press). [ค่า Impact Factor = 1.82]

ปีที่ 2 ลำดับที่ 2

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. Preattentive Auditory Processing of Vowel Duration Changes in Thai Words: Comparison between Native- and Non-Speakers of Thai. *Brain Topography* 2005 (in press). [ค่า Impact Factor = 1.82]

ปีที่ 2 ลำดับที่ 3

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. Preattentive Processing of Cluster vs. Non-cluster Initial Consonants in Thai. *Brain Topography* 2005 (in press). [ค่า Impact Factor = 1.82]

International Conference Contributions

ปีที่ 1 ลำดับที่ 1

Sittiprapaporn, W., Chindaduangratn, C., Kotchabhakdi, N. (2004). Auditory Preattentive Processing and Brain Lateralization of Long Vowel in Monosyllabic Thai Words. *Proceeding of the 1st KMITL International Conference on Integration of Science and Technology for Sustainable Development. 1:7-10.*

ปีที่ 2 ลำดับที่ 1

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Auditory Preattentive Processing of Vowel Duration Changes in Thai Words: Comparison between Native and Non-Native Speakers of Thai, Poster presented at *the Neurosciences and Music II,* organized by the Fondazine Pierfranco e Luisa Mariani ONLUS, Leipzig, Germany, May 2005.

ปีที่ 2 ลำดับที่ 2

Sittiprapaporn, W., Chindaduangratn, C. and Kotchabhakdi, N. (2005) The Perception of Short vs. Long Vowels in Monosyllabic Thai Words as Indexed by Mismatch negativity Brain Responses, Poster presented at the *X. International Congress for the Study of Child Language,* organized by the Congress Organization Thomas Wiese GmbH, Berlin, Germany, July 2005.

ปีที่ 2 ลำดับที่ 3

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Preattentive Processing of Cluster vs. Non-cluster Initial Consonants in Thai, Poster will be presented at *the 16th Meeting of the International Society for Brain Electromagnetic Topography (ISBET2005),* organized by the International Society for Brain Electromagnetic Topography, Bern, Switzerland, October 5 -8, 2005.

ปีที่ 2 ลำดับที่ 4

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Preattentive Auditory Processing of Vowel Duration Changes in Thai Words: Comparison between Native- and Non-Speakers of Thai, Poster will be presented at *the 16th Meeting of the International Society for Brain Electromagnetic Topography (ISBET2005),* organized by the International Society for Brain Electromagnetic Topography, Bern, Switzerland, October 5 -8, 2005.

ปีที่ 2 ลำดับที่ 5

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Source Generator of Across-and Within-Category Changes of Vowel in Consonant-Vowel (CV) Syllable: Low-resolution Electromagnetic Tomography (LORETA) Study, Poster will be presented at *the 16th Meeting of the International Society for Brain Electromagnetic Topography (ISBET2005),* organized by the International Society for Brain Electromagnetic Topography, Bern, Switzerland, October 5 -8, 2005.

10. ระเบียบวิธีในการวิจัย

อาสาสมัครเป็นกลุ่มคนไทยที่มีความสามารถเข้าใจภาษาไทย (กลาง) ได้ทั้งในระดับการฟัง พูด อ่าน และเขียน และกลุ่มชาวต่างชาติที่ไม่มีความสามารถเข้าใจภาษาไทย (กลาง) ได้ทั้งในระดับการฟัง พูด อ่าน และเขียน ทั้งเพศชายและเพศหญิง โดยมีช่วงอายุระหว่าง 18 ถึง 40 ปี รวมทั้งหมด 120 คน แบ่งเป็น ชาวไทย 65 คน และชาวต่างชาติ 55 คน ซึ่งอาสาสมัครทั้งหมดเป็นผู้ที่มีสุขภาพดี สมบูรณ์แข็งแรง และไม่เป็นโรคเกี่ยวกับระบบประสาทและไม่ได้ใช้ยาใดๆในระหว่างวิจัย ตลอดจนมีความสามารถในการได้ยินที่ปรกติและผ่านการทดสอบความสามารถทางด้านภาษา รวมทั้งการทดสอบความถนัดในการใช้มือ (Edinburgh Handedness Inventory) เพื่อให้แน่ใจว่า อาสาสมัครมีความชัดเจนในเรื่องของการถนัดใช้มือข้างขวามากกว่าข้างซ้ายหลังจากนั้น อาสาสมัครได้รับฟังคำอธิบายเกี่ยวกับรายละเอียดต่างๆของการทดลองที่อาสาสมัครมีส่วนร่วม พร้อมทั้งได้รับการตอบคำถามจากคณะผู้วิจัยจนกระจ่างชัดโดยไม่ได้ปิดบังซ่อนเร้น และเซ็นชื่อในหนังสือยินยอม (Informed Consent Form) ในการทำวิจัยอย่างเต็มใจในการศึกษาวิจัย อาสาสมัครจะได้รับฟังในเสียงพูดภาษาไทยที่เป็นคำพยางค์เดียว (Monosyllabic word) ซึ่งประกอบด้วย 4 กลุ่มคำ ดังนี้

กลุ่มคำที่ 1 คือ กลุ่มคำโครงสร้างพยางค์เดียวที่มีหน่วยเสียงสระที่ต่างประเภทกัน ได้แก่ คำว่า “ปอ-ปี และ ปอ-โป” เป็นต้น

กลุ่มคำที่ 2 คือ กลุ่มคำโครงสร้างพยางค์เดียวที่มีความสั้นยาวของหน่วยเสียงสระที่แตกต่างกัน ได้แก่ คำว่า “คาม-คำ และ ขำ-ขาม” เป็นต้น

กลุ่มคำที่ 3 คือ กลุ่มคำโครงสร้างพยางค์เดียวที่มีหน่วยเสียงวรรณยุกต์ที่แตกต่างกัน ได้แก่ คำว่า “ข้าม-ขำ และ ขาม-ข้าม” เป็นต้น

กลุ่มคำที่ 4 คือกลุ่มคำโครงสร้างพยางค์เดียวที่มีหน่วยเสียงพยัญชนะต้นเสียงควบกล้ำและไม่ควบกล้ำ ได้แก่ คำว่า “กาง-กลาง” เป็นต้น (ดูรายละเอียดเพิ่มเติมจากผลการศึกษา) ซึ่งอาสาสมัครจะดังยินเสียงผ่านทางหูฟังในระดับเสียงพูดปรกติซ้ำๆกันทุกๆ 1.24 นาที สลับกันแบบสุ่ม (Randomly Repetition) โดยมีเงื่อนไขว่าอาสาสมัครจะต้องไม่สนใจเสียงพูดดังกล่าวโดยการเบี่ยงเบนความสนใจด้วยการดูภาพยนตร์ ขณะเดียวกันอาสาสมัครจะได้รับการวัดคลื่นสมองในขณะที่ดูภาพยนตร์และไม่สนใจต่อเสียงพูดดังกล่าว โดยเวลาในการวัดคลื่นสมองประมาณ 1.5 ถึง 2 ชั่วโมง ส่วนข้อมูลที่อยู่ในรูปของคลื่นสมองของอาสาสมัครทั้งหมดจะนำไปวิเคราะห์ และแสดงผลในรูปของภาพการทำงานของสมอง (Neuroimaging)

จากการทบทวนวรรณกรรมที่เกี่ยวข้องในการศึกษาวิจัยทางด้านการทำงานของสมองในการจำแนกความแตกต่างของหน่วยเสียงในภาษานั้น คลื่นของสมองที่เรียกว่า มีสแมทเนกาทิวิตี้ (Mismatch Negativity, MMN) ซึ่งเป็นคลื่นของสมองที่ได้จากการจำแนกเสียงต่างๆทั้งที่เป็นเสียงในภาษาและไม่เป็นเสียงในภาษาในขณะที่ไม่ตั้งใจ (Unattended Auditory Perception) ซึ่งมีสแมทเนกาทิวิตี้ดังกล่าว ได้ถูกนำมาใช้กันอย่างแพร่หลายโดยเฉพาะอย่างยิ่งนักประสาทวิทยาศาสตร์ที่มุ่งศึกษาในด้านการทำงานของสมองที่เกี่ยวข้องกับภาษา (Neurolinguistics) โดยส่วนใหญ่จะมุ่งเน้นศึกษาเฉพาะกลุ่มของภาษาที่ไม่ใช้เสียงวรรณยุกต์ในการ

กำหนดความหมายของคำ อาทิเช่น ภาษาอังกฤษ เป็นต้น ซึ่งปัจจุบัน ยังไม่ปรากฏว่ามีการนำเอา

มีสแมทเนกาทิวิตี้ มาใช้ในการศึกษาถึงการทำงานของสมองของกลุ่มผู้ที่ใช้ภาษาที่ใช้วรรณยุกย์ในการสื่อสารแต่อย่างใด ดังนั้น การศึกษาวิจัยฉบับนี้จึงได้ทำการศึกษาว่า จะสามารถนำเอามีสแมชเนกาทิวิตี้ดังกล่าว มาใช้ในการศึกษาการทำงานสมองของอาสาสมัครทั้งชาวไทยและชาวต่างชาติ ในการจำแนกระบบคำในภาษาไทย ในการศึกษาครั้งนี้ ผู้วิจัยได้ใช้หน่วยเสียงต่าง ๆที่มีความแตกต่างกันทั้งทางด้านพยัญชนะ สระ และวรรณยุกต์ ที่ประกอบกันในโครงสร้างของระบบคำในภาษาไทย ซึ่งสามารถแบ่งออกเป็น 3 กลุ่ม ตามวัตถุประสงค์ของการศึกษา ดังนี้ (1) การศึกษาที่เกี่ยวกับการจำแนกประเภทของสระ ประกอบด้วย (ก) คำว่า /ปี/ มีเสียงสระส่วนหน้า ไม่ห่อปากและวรรณยุกต์เสียงกลางระดับ (ข) คำว่า /โป/ มีเสียงสระส่วนหน้า ห่อปากและวรรณยุกต์เสียงกลางระดับ และ (ค) คำว่า /ปอ/ มีเสียงสระส่วนหลัง ห่อปากและวรรณยุกต์เสียงกลางระดับ (2) การศึกษาที่เกี่ยวกับการจำแนกความแตกต่างทางด้านเสียงวรรณยุกต์และความยาวของสระ ประกอบด้วย (ก) คำว่า /คาม/ มีวรรณ์ยุกต์เสียงกลางระดับ (เสียงสามัญ) และสระเสียงยาว (ข) คำว่า /ข้าม/ มีวรรณยุกต์เสียงสูงตก (เสียงโท) และสระเสียงยาว (ค) คำว่า /ขาม/ มีวรรณยุกต์เสียงต่ำขึ้น (เสียงจัตวา) และสระเสียงยาว (ง) คำว่า /คำ/ มีวรรณยุกต์เสียงกลางระดับ (เสียงสามัญ) และสระเสียงสั้น (จ) คำว่า /ค่ำ/ มีวรรณยุกต์เสียงสูงตก (เสียงโท) และสระเสียงสั้น (ฉ) คำว่า /ขำ/ มีวรรณยุกต์เสียงต่ำขึ้น (เสียงจัตวา) และสระเสียงสั้น และ (3) การศึกษาที่เกี่ยวกับการจำแนกประเภทของพยัญชนะชนิดควบกล้ำและไม่ควบกล้ำ ประกอบด้วย (ก) คำว่า /กาง/ มีเสียงพยัญชนะชนิดไม่ควบกล้ำและมีวรรณยุกต์เสียงกลางระดับ (เสียงสามัญ) และ (ข) คำว่า /กลาง/ มีเสียงพยัญชนะชนิดควบกล้ำและมีวรรณยุกต์เสียงกลางระดับ (เสียงสามัญ)

ในการศึกษาครั้งนี้ ผู้วิจัยได้ทำการตรวจวัดการทำงานของสมองในรูปแบบของคลื่นสมองจากการตรวจวัดสัญญาณจากอิเล็คโทดจำนวน 21 อิเล็คโทรด ที่ได้ติดตามบริเวณต่างๆ ของสมอง จากการศึกษาพบว่าสมองมีการทำงานที่เด่นทั้งสองซีกในด้านต่างๆดังต่อไปนี้

(1) การจำแนกสระที่ต่างประเภทกัน เช่น สระส่วนหน้า ไม่ห่อปาก และสระส่วนหลังห่อปาก การจำแนกสระเสียงสั้นออกจากสระเสียงยาว

(2) การจำแนกการเปลี่ยนแปลงของสระจากสระเสียงยาวไปเป็นสระเสียงสั้นได้ดีกว่าการเปลี่ยนจากสระเสียงสั้นไปเสียงยาว

(3) การจำแนกการเปลี่ยนแปลงของเสียงวรรณยุกย์จากเสียงระดับไปเป็นเสียงต่างระดับ เช่น จากเสียงวรรณยุกต์กลางระดับ (เสียงสามัญ) ไปเป็นเสียงสูงตก (เสียงโท) แงะเสียงต่ำขึ้น (เสียงจัตวา)

(4) การจำแนกพยัญชนะเสียงควบกล้ำออกจากพยัญชนะเสียงไม่ควบกล้ำ

ทั้งนี้ ความสามารถในการทำงานของสมองทั้งสองซีกดังกล่าวสมารถพบได้ในอาสาสมัครทั้งสองกลุ่ม โดยเฉพาะอย่างยิ่ง คลื่นของสมองของอาสาสมัครชาวไทยจะเด่นกว่าคลื่นของสมองของชาวต่างชาติ ดังนั้น จากผลการศึกษาทำให้ทราบว่า มีสแมทเนกาทิวิตี้ สามารถนำมาใช้ในการตรวจวัดการทำงานของสมองในระบบคำในภาษาไทยด้านต่างๆดังต่อไปนี้ได้ คือ (ก) การจำแนกประเภทของสระ (ข) การจำแนกความแตกต่างทางด้านความสั้นยาวของสระ (ค) การจำแนกความแตกต่างทางด้านวรรณยุกต์ และ (ง) การจำแนกประเภทของพยัญชนะชนิดควบกล้ำและไม่ควบกล้ำ เป็นต้น ทั้งนี้ ความสามารถของสมองดังกล่าวจะใช้เวลาในการจำแนกเสียงตั้งแต่ระยะเวลา 2.5 ถึง 4.2 วินาที (150-250 msec) ตามแต่ละประเภทของคำ ดังนั้น จึงอาจกล่าวได้ว่า ค่ามีสแมทเนกาทิวิตี้ ที่ได้จากการจำแนกระบบคำในภาษาไทยสามารถนำมาใช้เป็นเครื่องชี้วัดความสามารถของสมองของอาสาสมัครในการเรียนรู้ภาษาได้ทั้งในระดับที่ตั้งใจ

และไม่ตั้งใจ (Attended and Unattended Auditory Processing) –v’สมองทั้งสองซีก ทั้งนี้กลไกการทำงานแบบอัตมัติดังกล่าว จึงสามารถนำไปใช้เป็นหลักเกณฑ์สากลใน

การจำแนกเสียงในภาษาของทั้งที่ใช้วรรณยุกต์และไม่ใช้วรรณยุกต์ในการกำหนดความหมายของคำ

ORIGINAL STUDIES

Contributions to Peer-Reviews Journals

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C., and Kotchabhakdi, N. Preattentive Discrimination of Across-category and Within-category Change in Consonant-Vowel Syllable. ***NeuroReport*** 2005; 16(13): 1513-1518.

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C. and Kotchabhakdi, N. Electrophysiological Measurement of Vowel Duration Changes in Monosyllabic Thai Words. ***NeuroReport*** 2005 (Submitted).

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C. and Kotchabhakdi, N. Auditory Preattentive Processing of Cluster vs. Noncluster Initial Consonant Perception in Monosyllabic Thai Words: Preliminary Study. ***NeuroReport*** 2005 (Submitted).

Sittiprapaporn, W., Tervaniemi, M., Chindaduangratn, C. and Kotchabhakdi, N. Functional Specialization of the Human Auditory Cortex for Processing of Lexicality in Thai Language ***Cognitive Brain Research*** (In Preparation).

Published Abstracts

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. Source Generator of Across-and Within-Category Changes of Vowel in Consonant-Vowel (CV) Syllable: Low-resolution Electromagnetic Tomography (LORETA) Study**. *Brain Topography*** 2005 (in press).

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. Preattentive Auditory Processing of Vowel Duration Changes in Thai Words: Comparison between Native- and Non-Speakers of Thai. ***Brain Topography*** 2005 (in press).

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. Preattentive Processing of Cluster vs. Non-cluster Initial Consonants in Thai. ***Brain Topography*** 2005 (in press).

Awikunparsert, C., Sittiprapaporn, W., Chindaduangratn, C., and Kotchabhakdi, N. Neurophysiological Evidence of Memory Traces for Words with Vowel of Lexical Tones. ***International Journal of Behavioral Medicine*** 2004; 11:107.

Sittiprapaporn, W., Awikunparsert, C., Chindaduangratn, C., and Kotchabhakdi, N. Auditory Processing for Mismatch Negativity Responses to Across-and Within-category Change of Vowels. ***International Journal of Behavioral Medicine*** 2004; 11:111.

Sittiprapaporn, W., Awikunparsert, C., Chindaduangratn, C., and Kotchabhakdi, N. Memory Traces for Cluster vs. Noncluster Consonants as Shown by Mismatch Negativity. ***International Journal of Behavioral Medicine*** 2004; 11:111.

International Conference Contributions

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Source Generator of Across-and Within-Category Changes of Vowel in Consonant-Vowel (CV) Syllable: Low-resolution Electromagnetic Tomography (LORETA) Study, Poster will be presented at ***the 16th Meeting of the International Society for Brain Electromagnetic Topography (ISBET2005),*** organized by the International Society for Brain Electromagnetic Topography, Bern, Switzerland, October 5 -8, 2005.

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Preattentive Auditory Processing of Vowel Duration Changes in Thai Words: Comparison between Native- and Non-Speakers of Thai, Poster will be presented at ***the 16th Meeting of the International Society for Brain Electromagnetic Topography (ISBET2005),*** organized by the International Society for Brain Electromagnetic Topography, Bern, Switzerland, October 5 -8, 2005.

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Preattentive Processing of Cluster vs. Non-cluster Initial Consonants in Thai, Poster will be presented at ***the 16th Meeting of the International Society for Brain Electromagnetic Topography (ISBET2005),*** organized by the International Society for Brain Electromagnetic Topography, Bern, Switzerland, October 5 -8, 2005.

Sittiprapaporn, W., Chindaduangratn, C. and Kotchabhakdi, N. (2005) The Perception of Short vs. Long Vowels in Monosyllabic Thai Words as Indexed by Mismatch negativity Brain Responses, Poster presented at the ***X. International Congress for the Study of Child Language,*** organized by the Congress Organization Thomas Wiese GmbH, Berlin, Germany, July 2005.

Sittiprapaporn, W., Khampan, W., Wang, T-C., Ananchaipatthana, P., Chindaduangratn, C. and Kotchabhakdi, N. (2005) Auditory Preattentive Processing of Vowel Duration Changes in Thai Words: Comparison between Native and Non-Native Speakers of Thai, Poster presented at ***the Neurosciences and Music II,*** organized by the Fondazine Pierfranco e Luisa Mariani ONLUS, Leipzig, Germany, May 2005.

Sittiprapaporn, W., Chindaduangratn, C., Kotchabhakdi, N. (2004). Auditory Preattentive Processing and Brain Lateralization of Long Vowel in Monosyllabic Thai Words. ***Proceeding of the 1st KMITL International Conference on Integration of Science and Technology for Sustainable Development. 1:7-10.***

The studies are also referred to in the text.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT.)

TABLE OF CONTENTS (CONT.)

LIST OF TABLES

LIST OF TABLES (CONT.)

LIST OF FIGURES

LIST OF FIGURES (CONT.)

LIST OF FIGURES (CONT.)

Figure Page

LIST OF FIGURES (CONT.)

Figure Page

LIST OF FIGURES (CONT.)

LIST OF ABBREVIATIONS

ACPs = Association Cortex Potentials

AEPs = Auditory Evoked Potentials

ANOVA = One Way Analysis of Variance

BA = Brodmann’s Area

CBF = Cerebral Blood Flow

CECIL = Computerized Extraction of Components of Intonation in Language

CI-500 = Computer Interface 500

CSR = Central Sound Representation

CV = Consonant-Vowel Syllable Structure

D.S. = Digit Span

DSB = Digit Span Backward

DSF = Digit Span Forward

DST = Digit Span Total

EEG = Electroencephalogram

ERPs = Event-Related Potentials

FFT = Fast Fourier Transformation

fMRI = Functional Magnetic Resonance Imaging

Fo = Fundamental Frequency

Fz = Frontal Electrode Site

GFP = Global Field Power

Hz = Hertz

LIST OF ABBREVIATIONS (CONT.)

ISI = Inter-stimulus Interval

L1 = First Language

L2 = Second Language

LH = Left Hemisphere

LORETA = Low Resolution Electromagnetic Tomography

MEG = Magnetoencephalography

MMN = Mismatch Negativity

MMNm = Magnetic Mismatch Negativity

MRI = Magnetic Resonance Imaging

NTS = Native Thai Speaker

N100 = Negative 100

P300 = Positive 300

PET = Positron Emission Tomography

RH = Right Hemisphere

SOA = Stimulus onset Asynchrony

SP Maps = Structure Probability Maps

St. = Semitones

TD = Talairach Daemon

UTT = Utterance

VOT = Voice on Set Time

WAV = Wave File

CHAPTER I
INTRODUCTION

This chapter outlines some of the features of speech perception and the human brain and point out the advantages of using the passive oddball paradigm in the electroneurophysiological study of speech perception. It is then briefly reviewed the neurophysiology and functional neurophysiology of speech perception. In these reviews, particular attention is paid to the roles of hemisphere in the speech perception. Finally, the importances of neurophysiological features of mismatch negativity (MMN), which are relevant to this study, are also outlined.

1.1. Overall Views

The study of speech in the present day has proceeded to the level of understanding the neural mechanisms underlying different aspects of speech and their localization of the brain. The research in the past has shown that both the production and perception of human speech involved specialized neural devices. The primary focus has been on hemispheric dichotomies in the processing of speech perception and the major focus has been shifted to the development of neural networks subserving speech perception. Two major, competing lines of investigation have emerged concerning hemispheric specialization for speech perception. One emphasizes task-dependent or domain-specific effects (Van Lanker, 1980; Ross and Mesulam, 1979), and the other cue-dependent effects that cut across task domains (Robin et al., 1990; Ivry and Roberson, 1998). In speech perception, task-specific hypotheses assume that unique, neural mechanisms are recruited for the speech domain, whereas cue-dependent hypotheses claim that speech processing is subserved by neurobiological mechanisms specialized for particular aspects of the acoustic signal, irrespective of communicative or linguistic relevance. However, these hypotheses has been against from some evidences of neurological impaired populations with unilateral lesions to the left (LH) or right hemisphere (RH) as well as from some studies using dichotic listening procedures or the Wada technique

(Gandour, 2000). Within the past decade, empirical evidence has also begun to appear from both lesion-deficit and functional neuroimaging studies. Lesion studies tell us what areas are necessary for normal functioning, while functional neuroimaging studies tell us what areas participate in that function. The functional neuroimaging also give us a window on language processing *in vivo* in the normal functioning human brain (Gandour, 2000). However, the physiological method of describing the anatomy of human speech still helps us to indicate what elements have a functional value with regards to speech and what anatomical elements are irrelevant.

When engaged in a conversation, listeners tune in to the relevant stream of speech and filter out irrelevant speech input that may be present in the same environment. Nonetheless, attention might be involuntarily diverted to meaningful items coming from an ignored stream, like in the well-known own-name effect (Moray, 1959). This brings up the question of to what extent speech is processed in the ignored streams. In auditory perception, the occurrence of the deviant (infrequent) stimulus after a sequence of the standard (Frequent) stimuli tends to elicit mismatch negativity (MMN) in event-related potentials (ERP) and its magnetic equivalent called the magnetic mismatch field (MMF) in magnetoencephalography (MEG). The MMN/MMF component may be considered to reflect the pre-attentive auditory memory processes and represents neuronal correlates of change detection and sound discrimination (Näätänen, 2001).

Previous studies showed that for sinusoidal tones, the MMF is sensitive to the direction of a change within the stimulus, being more robustly activated for duration shortening or pitch falling as opposed to lengthening or leveling (Inouchi et al., 2002). These studies also revealed no significant differences between subjects who spoke a pitch-accent language (Japanese) and those who did not (English). It has been reported that MMN/MMF is indeed sensitive to cross-linguistic relevance. Unlike short-to-long vowel duration and falling-to-level pitch changes, long-to-short duration and level-to-falling pitch changes elicited a prominent MMF bilaterally for both groups, peaking at around 100 ms after change onset for duration and 200 ms for pitch. The MMF component is sensitive is sensitive to vowel shortening rather than lengthening and to pitch falling rather than leveling (Inouchi et al., 2002, 2003).

Using Thai words, a lexical tone in a tonal language, the current study investigated whether there would be an influence of cross-linguistic relevance on the elicitation of the MMN in response to Thai words for both native and nonnative speakers of Thai. The present study is thus concerned with an issue related to this question.

1.2. General Background of Human Speech

Human communication is made up of three domains: speech, language, and pragmatics, which can be fully specified as a set of rules, units, and practices. Prosodic function spans these domains. Speech and language are largely lateralized to the left cerebral hemisphere, while pragmatic competence appears to require an intact right hemisphere. The term "speech" is best used to refer to motor output and perceptual skills; "language" covers internal, mental knowledge; "pragmatics" expands the scope of study to overall, actual language use. Prosodic competence spans these domains. Prosody, or melody of speech, being made up of pitch, duration, intensity, and voice quality, occurs in the domain of speech, and in the capacity, can be quantified to a considerable degree. Prosodic meanings can be linguistic, as in the grammatical contrast between statement and question; or paralinguistic, as in the communication of attributes, emotional meanings and pragmatic factors. Prosody is receiving increasing attention in unimpaired subjects (Banse and Scherer, 1996; Crystal, 1696; Scherer, 1986; Freese and Maynard, 1998) and subjects with brain damage. Prosody cues linguistic, attitudinal, emotional, and pragmatic information, but it has been most thoroughly studied in its role of signaling emotional meanings (Breitenstein et al., 1997).

1.3. Physiological Basis of Human Speech

It is generally accepted that speech is an acoustic signal comprised of multiple co-occurring frequencies, called formats. The communication through the medium of language is based on four major constituents: lexicon (vocabulary), syntax (grammar), *prosody*, and kinesics. The smallest articulated feature of language is known as a segment, which is closely aligned to the syllable (Kent and Read, 1992; Ledefoeged, 1975; Ross, 1999). Segments are also known as the primary phonetic building blocks for words, which form the lexicon (Ross, 1999). Theoretically, in autosegmental phonology (Goldsmith, 1979, 1990; Halle and Vergnaud, 1980), a phonological representation consists of several layers or *tier*. The axis of this representation consists of a sequence of timing slots. Each slot contains phonetic information. This sequence of slots is called the *skeletal tier.* Such slots are therefore indicated by C and V for consonant and vowel, respectively. The slots at the skeletal tier are normally grouped into successive syllables. These are then represented at the syllable tier. Such strings of C and V slots within a syllable can be further organized into syllable constituents. The phonetic content of the slots is so-called both qualitative and prosodic. Qualitative features represent various articulatory gestures to be made in producing a consonant or vowel. They are specified at the segment tier. Current phonological theories posit a separate prosodic tier that specifies metrical structure (Levelt 1989, Leberman and Prince 1977, Selkirk 1984, Baum and Pell 1999).

Prosody and kinesics generally constitute the paralinguistic elements of language and play an equally prominent role on the organization of human communication and discourse (Ross, 1999; Bolinger, 1980). *Prosody* refers to the distribution of stress and melodic contour in speech. Modulation of prosody can be used to impart affective tone, introduce subtle grades of meaning, and vary emphasis in spoken language. Normal speech is always subject to the modulation of prosody. These modulations may be minimal, signaling subtle connotation, or, more significant, changing the same sentence from a statement to a question (Mateer, 1989). Prosody, therefore, appears to contribute a great deal to the overall effectiveness and appropriateness of oral communication (Mateer, 1989). In general, prosody is a mode of communication, which provides a parallel channel to speech. Prosodic features, unlike other linguistic components, are often produced without conscious intention and are open to forms of interpretation, which rely on emotional, non-cognitive processes. The communicative content of many prosodic signals parallels that of stereotypic call vocalizations characteristic of communication systems of subhuman species (Dogil et al., 2002). It has often been argued that the neuroanatomical basis for these call vocalizations should be fundamentally different from that of the symbolic aspects of human communication (Dogil et al., 2001; see also Deacon, 1984). But unlike calls of other species, prosodic organization of human communication is continuous and highly correlated with the semantic, syntactic, morphological and segmental organization of speech, which is fully under neocortical control. Regardless of function, there exist only three prosodically active phonetic parameters: duration, intensity and pitch (Dogil et al., 2002). According to the neurobiology of speech prosody initiated by Monrad-Krohn, the speech prosody can be divided into several types' intrinsic (linguistic) prosody, dialectal and idiosyncratic prosody, intellectual prosody, emotional (affective) prosody, and inarticulate prosody.

Theoretically, prosody is known as a suprasegmental feature of language that conveys information beyond that transmitted by word choice and word order alone (Kent and Read, 1992; Ross, 1999; Bolinger, 1980; Crystal, 1975; Monrad-Krohn, 1963; Van Lancker, 1980). Prosodic features of speech include acoustical properties: timing, intensity, and fundamental frequency (*Fo*) (Gandour et al., 1989). Variations in these three acoustical properties of the speech waveform include duration, amplitude, and fundamental frequency (*Fo*). The principal prosodic or suprasegmental features of speech which correlate of these attributes thus include tone, intonation, stress, length, and rhythm (Lehiste, 1970, 1976; Ledefoged, 1982; Gandour et al., 1988). Prosodic features such as rhythm, pitch, pause etc., carry more substantial contribution to the intelligibility of speech (Darwin, 1975; Studdert-Kennedy, 1980; Kadota, 1987). Speech prosody is thus comprised of multiple phonetic correlates (e.g., pitch, duration, and loudness) that subserve multiple or plays a critical role in linguistic and paralinguistic functions (Gandour, 1998; Gandour et al., 2000*)*. The term *"prosodic feature"* may be regarded as abstractions apart from the consonant and vowel system (Firth, 1948, 1970). Such abstractions may be made at the syllable, word, or sentence level. *Syllable prosodies* include tone, quantity, and those properties, which mark the beginning or end of a syllable. *Word prosodies* include tonal and quantitative features, stress, and the means whereby syllable is linked with syllable. *Sentence prosodies* include sentence tone, and the means used to mark beginnings and ends of phrases and sentences, and to connect phrase with phrase or sentence with sentence (Henderson, 1949, 1970). The phonological structure of the sentence and the words, which comprise it, are to be expressed as a plurality of systems of interrelated phonematic and prosodic categories (Firth, 1948, 1970). The prosodies of the sentence and the word tend to be dominant emphasizing the phonetics and phonology of synthesis.

Speech prosody generally involves the “musical” aspects of language-variations in the pitch, length, and loudness of spoken utterances (Cruttenden, 1997; Lehiste, 1970). Like other prosodic features, falling and rising pitch patterns are said to be “suprasegmental” because they co-occur with strings of sound segments and thus constitute a phonetic overlay that is superimposed on the segmental speech stream (Snow, 2000). The principal linguistic correlates of these attributes include stress, intonation, tone, length, and rhythm. Because these acoustical properties systematically correspond to features of linguistic structure, we can acquire knowledge about the structure and processing of language by measuring these acoustical properties of the speech wave (Cooper and Sorensen, 1981). Speech prosody is known as a multifaceted process that requires integration of multiple cues, some of which appear to be processed in the left cerebral hemisphere, others in the right cerebral hemisphere. Changes in speech prosody can influence both linguistic and nonlinguistic aspects of the speech signal. Variations in these prosodic features provide auditory cues to phonological, lexical and grammatical unit (linguistic prosody), as well as emotional states (paralinguistic or affective prosody) (Gandour, 1998; Gandour et al., 1995, 2000; Hsieh et al., 2001). However, speech prosody may serve different functions; e.g., linguistic, and paralinguistic functions.

1.4. Language Processing in Neuroscience

1.4.1. Functional Neuroanatomy of Language and Speech

The study of the correlation between brain and language has always lagged behind the study of other aspects of behavior and cognition due to the lack of an animal model. The lack of a homologous to language in other species precludes the attempt to model language in animals, and our understanding of the neural basis of language must be pieced together from other sources (Dronkers et al., 2000). Thus, the scientific community predominantly had to rely on the evidence provided by the correlation of neurological disorders with language disorders, apart from selection brain recording and stimulation techniques performed during neurosurgery (Dogil et al., 2002). The most important source has been the study of language disorders, known as *aphasia*, which are caused by focal brain lesions that result from stroke or head injury. The early study of aphasia revealed that damage to each of two cortical areas, one in the lateral frontal region, the other in the posterior superior temporal lobe, was associated with a major and linguistically different profile of language impairment. The two cortical areas are *Broca's area* and *Wernicke's area.* Dogil *et al.* (2002) recently suggests the neuroanatomical structure underlying the production of speech and structure dependent syntactic operations. Both of these neural substrates of language can be represented as networks of focal areas widely distributed in the brain. The speech network consists of *Motor Cortex—Supplementary Motor Area—Cerebellum—Anterior Insula* of the dominant hemisphere. A small section of the insula is important for planning or coordinating the articulatory movements necessary for speech. Patients who have lesions in this area have difficulty pronouncing phonemes in their proper order; they usually combinations of sounds that are very close to the great word. However, these patients have no difficulty in perceiving speech sounds or recognizing their own errors and do not have trouble in finding the word, only in producing it (Dronker et al., 2000). Concerning language production the processing model of Levelt (1989, 1999) is often used in several studies. Not only is it the model with the strongest experimental support, but it is also the one that understands speech and language production as a modular process, including several self-contained tasks which can be easily tested using the image-subtraction methodology (Dogil et al., 2002). Levelt's "blueprint' for the speaker help to find the correlates for those components in the working brain. The Levelt's blueprint model distinguishes three building blocks in the phonological/phonetics system: morpho-phonological encoding, phonetic encoding and articulation. The neural correlates of these three components concern the role of address frames in prosody generation.

The study of language and the brain concerns the delineation of the neural substrate underlying the perception and production of prosody in relation to both linguistic and paralinguistic (affective) components of the speech signal. In addition to the relationship between prosody and speech perception, Lehiste (1973) stressed the importance of speech rhythm among several prosodic variables in such a way that it might play an essential part in constructing a processing unit of speech. Furthermore, researchers such as Martin (1972), Lehiste (1979), and Donovan and Darwin (1979) advocated a top-down, prosody-oriented approach to speech perception and production, in which some prosodic characteristics are claimed to constitute a fundamental framework for the subsequent segmental processing and production (Kadota, 1987). Speech prosody has become an increasingly useful medium for studying the role of the left and right cerebral hemispheres in the control of acoustical properties in both linguistic and emotional domains (Ledefoged, 1982). Whereas a number of experimental studies have suggested that the right cerebral hemisphere is generally dominant for emotional processing (Ross, 1981; Ross and Mesulam, 1979; Bryden and Ley, 1983; Seiberman and Weingatner, 1986), and therefore for affective prosody, further research has implicated right cerebral hemisphere involvement for linguistic prosody as well (Behrens, 1989; Shapiro and Danly, 1985; Weintraub et al., 1981, Ryalls and Behrens, 1988). Gandour and colleagues (1988) has proposed another view on lateralization of prosodic functions that linguistic and affective prosody may be generated by the same phonetic mechanisms, but that temporal and spectral aspects are hemispherical specialized. The right cerebral hemisphere processes Fo information but not other prosodic features such as timing or intensity (Gandour et al., 1995; Robin et al., 1990, 1991). Left cerebral hemisphere contributions to prosody may be related to the processing of timing information (Gandour, 1998). According to Hughes, et al. (1983), the right cerebral hemisphere appears to exert a major influence in the organization of these paralinguistic functions. Consequently, the modulation of prosodic elements may precede the emergence of phonetic segmentation during the early acquisition and development of human language (Crystal, 1973; Lewis, 1936).

In addition, speech prosody has been encapsulated in the phonological component of psycholinguistic processing models of both speech perception (Levell, 1999), and production (Cutler and Chen, 1997). Prosodic information is important in prelexical processes of phonological encoding, and has been shown to constrain word activation in tonal (Cutler and Chen, 1997), pitch accent (Cutler and Otake, 1999), and lexical stress languages (Mattys, 2000). It is, however, difficult to identify any specific brain areas corresponding to the prelexical phonological processing state with functional neuroimaging because speech input generally triggers lexico-semantic processing (Fry, 1968). However, it is necessary for candidate brain areas to be demonstrated the insensibility to the lexical status of the stimuli, yet respond to phonemes, illegal non-words, pseudo-words, or words (Price et al., 1999). Because of its multiple functions (e.g., linguistic, attitudinal, affective) and its multiple phonetic cues (e.g., pitch, duration, loudness), the functional organization of the human brain, with respect to the specific localization of cortical regions involved in processing different aspects of speech prosody including the neural substrate of speech prosody, remains elusive and controversial as well as a fundamental question that has yet to be clearly answered.

1.4.2. Functional Neuroanatomy of Speech Perception

Language organization in the brain has been studied with positron emission tomography (PET) (Frith et al., 1991; Leblanc et al., 1992; Haxby et al., 1991; Petersen et al., 1993; Wise et al., 1991; Klein et al., 1994, 1995), intraoperative stimulation mapping (Ojemann et al., 1990, 1983, 1987, 1978; Schaffler et al., 1993), and functional MR imaging (Hinke et al., 1993; Yetkin et al., 1995; Binder et al., 1994; Shaywitz et al., 1994). PET has been shown activation in the left prefrontal cortex (Broca's area) caused by word generation or counting tasks (Wise et al., 1991). However, the cerebral organization of bilinguals with respect to language is still disputed, despite many investigations with different modalities (Chee et al., 1999).

Due to the limitations of the classical model of language localization, which fail to account for the functional overlap between Broca's and Wernicke's areas in producing and comprehending spoken language, as well as the involvement of other regions, both cortical and subcortical, a refined and extended model of language specialization is being developed to modify the classical Wernicke-Geschwind model (Dronkers, Pinker, and Damasio, 2000). Instead of a unidirectional pathway from posterior Wernicke's- to anterior (Broca's) regions, the arcade fusciculus provides for a bi-directional pathway between frontal and temporoparietal regions. Neuroimaging data shows that left frontal areas participate in language comprehension (Binder et al., 1996, 1997) and speech perception (Zatorre et al., 1992, 1996; Gandour et al., 2000, Hsieh et al., 2001). Left frontal regions mediate *perception* as well as production of both segmental and suprasegmental aspects of the speech signals. Involvement of left frontal regions in speech perception is compatible with the notion that articulatory recording of speech signals occurs during speech perception (Liberman and Mattingly, 1985, 1989). Executive functions (selective attention, inhibition, and active comparison) that mediate phonological processing appear to be a function of the rostal portion of the left pars triangularis (BA 45). Conversely, the left planum temporal and left posterior perisylvian cortex, i.e. secondary auditory cortex, have been implicated in a speech production task that masked out auditory input (Paulesu et al., 1996). These audio-vocal interactions in the auditory system are seen to be analogous to visuo-oculomotor interactions in the visual system. Converging evidence from lesion studies show that Broca's aphasics have comprehension problems in processing grammatical aspects of language (Berndt et al., 1996; Caplan et al., 1996), and that Wernicke's and conduction aphasics exhibit "subtle phonetic deficits" in speech production (Vijayan and Gandour, 1995).

Other cortical regions in the LH may be critically involved in language processing including prefrontal and cingulate areas that implement executive control and mediation of necessary memory and attentional processes (Dronkers et al., 2000). The rostral portion of the left pars triangularis may play an executive role in mediating between long-term phonological representations and a short-term phonological buffer (Hsieh et al., 2001). Consistent with previous studies on the attentional functions of anterior cingulate cortex and medial frontal gyrus (Cabeza and Nyburg, 1997; Carter et al., 1998).

The role of the insular in language remains somewhat controversial (Ardila et al., 1997; Ardila, 1999; Flynn et al., 1999). It is well-positioned anatomically to play a crucial role in neural networks involved in verbal communication. Damage to the insular has been proposed as one of the possible neural substrates of conduction aphasics (Damasio and Damasio, 1980), apraxia of speech (Dronkers, 1996), auditory agnosia (Habib et al., 1995), Dyslexia (Paulesu et al., 1996), and verbal memory (Manes et al., 1999). Previous neuroimging studies have implicated insular cortex in tasks involving articulation and speech coordination (Wise et al., 1999; Dronkers, 1996), short-term verbal memory (Paulesu et al., 1993), and music (Zatorre et al., 1994). Its precise function has been variously interpreted to be associated with coordination of speech articulation (Dronkers, 1996; Wise et al., 1999), subvocal rehearsal in verbal short-term memory (Paulesu et al., 1993), subvocal rehearsal in vocal pitch (Zatorre et al., 1994), and conversion of auditory-visual signals to phonological codes (Paulesu et al., 1996). The existence of possibly two parallel, separate pathways projecting from the temporo-parietal to the frontal language area, the first via the arcuate fasciculus, the second via the insular (Lichtheim, 1885; McCarthy and Warrington, 1984). Anatomically, insular cortex is connected to pre- and post-central opercula cortex; inferior frontal cortex; adjacent auditory cortex (BA 22, BA 41/42); several limbic structures, including anterior cingulate gyrus; as well as several thalamic nuclei (Mesulam and Mufson, 1985). Thus, it is well-positioned anatomically to play a crucial role in neural networks involved in verbal communication. The insular also appears to be a crucial element in the nonsemantic processing route. It is posited to be a link between *"unsegmented"* and parsed phonological codes (Paulesu et al., 1996). Its role is to covert a visual or auditory input signal into a parsed string of consonants and vowels *(segmental)*, and tones *(suprasegmental)* (Gandour et al., 2000).

Structures within the RH are proposed to be specialized for at least some aspects of pitch processing (Imaizumi et al., 1998; Robin et al., 1990; Zatorre et al., 1992, 1994). The right anterior insular as well as other IFG regions in the RH is possibly critical for production, covert as well as overt, of all complex volitional vocalizations, including vocal fundamental frequency and articulation (Perry et al., 1999). Consistent with a recent fMRI study of overt speaking and singing, reproduction of a non-lyrical tune elicited activation predominantly in the right motor cortex, the right anterior insular, and the left cerebellum, whereas the opposite response pattern emerged during the speech task (Riecker et al., 2000).

Only a few of language studies involving a crosslinguistic to determine the influence of linguistic experience on the perception of segmental (consonants, vowels) and suprasegmental (tones) information, reported to date with functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have included the native speakers of tone language and a nontone language (Gandour et al., 2000; Hseih et al., 2001), but not in bilingual volunteer subject. Positron Emission Tomography (PET) is used in a crosslinguistic study to compare pitch processing in a native speaker of two tone languages, Chinese and Thai, with those of English, a nontone language (Gandour et al., 1998, 2000). Participants from each language group are scanned under three active tasks (tone, pitch, and consonant) that required focused-attention, speeded-response, auditory discrimination judgments, and one passive baseline as silence. When comparing the tone to the pitch task, only the Thai group shows significant activation in the left frontal operculum. Activation of the left frontal operculum in the Thai group suggests that phonological processing of suprasegmental as well as segmental units occurs in the vicinity of Broca's area. Baseline subtractions, on the other hand, show significant activation in the anterior insula region for the English and Chinese groups, but not Thai, providing further support for the existence of possibly two parallel, separate pathways projecting from the temporo-parietal to the frontal language area (Gandour et al., 2000). However, Chinese and English listeners do not show the same LH lateralization as Thai listeners when making perceptual judgments of Thai tones. More generally, these differential patterns of brain activation across language groups and tasks support the view that pitch patterns are processed at higher cortical levels in a top-down manner according to their linguistic function in a particular language (Gandour et al., 2000). In the study of Hsieh and colleagues (2001), Chinese and English listeners are asked to make perceptual judgments of Chinese tones, consonants, and vowels (Hsieh et al., 2001). A crosslinguistic, positron emission tomography (PET) study is conducted to determine the influence of linguistic experience on the perception of segmental (consonants, vowels) and suprasegmental (tones) information. Chinese listeners show increased activity in left premotor cortex, pars opercularis, and pars triangularis across all tasks. English listeners, on the other hand, show increased activity in left inferior frontal gyrus regions only in the vowel task and in right inferior frontal gyrus regions in the pitch task. Findings suggest that functional circuits engaged in speech perception depend on linguistic experience. All linguistic information signaled by prosodic cues engages left hemisphere mechanisms. Storage and executive processes of working memory that are implicated in phonological processing are mediated in discrete regions of the left frontal lobe (Hseih et al., 2001).

1.5. Neurophysiological Features of Mismatch Negativity (MMN)

The mismatch negativity (MMN) (Näätänen et al., 1978) and its magnetic equivalent MMNm) is elicited by any discriminable change in some repetitive aspect of auditory stimulation, irrespective of the direction of the subject’s attention. This negative component of the auditory event-related potential (ERPs), usually peaking

100-300 ms from change onset (Sams et al., 1985), is based on, and reflects, neural traces by which the auditory cortex models the repetitive aspects of the acoustic past (Näätänen and Winkler 1999). These traces might contain sensory information on sound frequency, duration and inter-stimulus interval (ISI), but also on more complex aspects of auditory stimulation, such as rhythmic patterns or speech sounds (Näätänen and Winkler 1999). The properties of these traces (which usually last several seconds, although even permanent traces can be reflected (Näätänen and Winkler 1999)) can be probed by presenting infrequent deviant events in the sequence of repetitive events ('standard') (Näätänen and Winkler 1999). MMN is elicited even in the absence of attention, for example, in individuals in a coma a few days before the recovery of consciousness (Kane et al., 1993), which indicates that MMN indexes pre-attentive (attention-independent) auditory processing.

The main neural generators of MMN are bilaterally located in the supratemporal plane (Alho, 1995), which is indicated by dipole modeling (Scherg et al., 1989) and scalp current density map (Giard et al., 1990) of scalp-recorded event-related potentials, as well as by magnetic recordings (Levänen et al., 1996), intracranial MMN recordings in cats (Csepe et al., 1987), monkeys (Jaavitt et al., 1992) and humans (Halgren et al. 1995; Kropotov et al., 1995, 2000), and by positron emission tomography (Tervaniemi, et al., 2000), functional magnetic resonance imaging (Opitz et al., 1999; Celsis et al., 1999), and optical imaging data (Rinne et al., 1999). Furthermore, the exact locus of MMN in auditory cortex depends on the attribute (Giad et al., 1995) (and even on the complexity of stimulus configuration (Alho et al., 1996)) in which the change occurred. Therefore, one can conclude that the auditory processes that generate MMN originate, in the first place, in the auditory cortex. In addition, MMN also receives a contribution from a (mainly right hemispheric) frontal generator that appears to be triggered by this auditory-cortex change-detection process and be associated with the initiation of attention switch to the change (Escera et al., 2000).

The mismatch negativity (MMN) is a component of the auditory event related potential (ERP) which is elicited task-independently by an infrequent change in a repetitive sound. The MMN can be recorded in response to any discriminable change in the stimulus stream. The MMN data imply the existence of a sensory-memory trace in which the features of the frequently occurring standard stimuli are represented. One can probe this trace by presenting deviant stimuli of different magnitudes and thus indirectly determine the accuracy of this central sound representation. Several recent studies have shown that these representations govern attentive auditory discrimination ability in humans. Therefore the MMN is an objective, easily quantifiable index of the quality of sensory stimulus representations from which auditory percepts are built. The most recent studies have provided evidence that even complex, temporal, linguistic stimulus features and long-term learning effects are reflected in MMN responses, thus significantly broadening the theoretical scope of the MMN research. The MMN is consequently of great potential interest in attempts to understand central auditory function, its development, and various forms of its pathology.

Processing of sensory stimulus features is essential for humans in determining their responses and actions. If behaviorally relevant aspects of the environment are not correctly represented in the brain, then the organism's Behaviour cannot be appropriate. Without these representations our ability to understand spoken language, for example, would be seriously impaired. Cognitive neuroscience has consequently emphasized the importance of understanding brain mechanisms of sensory information processing, that is, the sensory prerequisites of cognition. Most of the data obtained, unfortunately, do not allow the objective measurement of the accuracy of these stimulus representations (Näätänen, 1992). In audition, recent cognitive neuroscience seems to have succeeded in extracting such a measure, however. This is the mismatch negativity (MMN), a component of the event-related potential (ERP), first reported by Näätänen, Gaillard, and Mäntysalo (1978). An in-depth review of MMN research can be found in Näätänen (1992) while other recent reviews also provide information on the generator mechanisms of MMN (Alho 1995), its magnetic counterpart, MMNm (Näätänen, Ilmoniemi and Alho, 1994), and its clinical applicability (Näätänen and Alho, 1995).

MMN is evoked by an infrequently presented stimulus ("deviant"), differing from the frequently-occurring stimuli ("standards") in one or several physical parameters like duration, intensity, or frequency (Näätänen, 1992). In addition, it is generated by a change in spectrally complex stimuli like phonemes, in synthesized instrumental tones, or in the spectral component of tone timbre. Also the temporal order reversals elicit an MMN when successive sound elements differ either in frequency, intensity, or duration The MMN is not elicited by stimuli with deviant stimulus parameters when they are presented without the intervening standards. The MMN has, therefore, been suggested to reflect change detection when a memory trace representing the constant standard stimulus and the neural code of the stimulus with deviant parameter(s) are discrepant.

The search for an objective index of change detection in the human brain can be traced back to 1975, with the proposition that stimulus deviation per se (irrespective of, e.g., stimulus significance, attentional mechanisms) should produce a measurable brain response (Näätänen, 1975). Experimental evidence for this suggestion was obtained in experiments conducted by Näätänen, Gaillard, and Mä ntysalo in 1975 (subsequently reported in 1978). In this dichotic listening study, the subject's task was to detect occasional deviant stimuli in the stimulus sequence presented to a designated ear while ignoring the concurrent sequence presented to the opposite ear. The irrelevant stimulus sequence included deviant stimuli that were physically equivalent to the deviant stimuli (targets) of the attended input sequence. The deviant stimuli were either tones of a slightly higher frequency or tones of a slightly greater intensity than the standard tones.

The deviant stimuli both in the attended and unattended stimulus sequence elicited negativity in the 100-200 ms latency range, which could not be seen in response to the standard stimuli. This negativity, usually described by the deviant-minus-standard difference wave, was very similar for the attended and ignored input sequences, suggesting that attention was not required. Näätänen et al. (1978) proposed that it may well be that a physiological mismatch process caused by a sensory input deviating from the memory trace 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 subject, perhaps even unmodified by the latter. On the basis of the relatively large MMN amplitudes above the temporal areas, the authors further suggested that the mismatch negativity reflects specific auditory stimulus discrimination processes taking place in the auditory primary and association areas... The latter processes are suggested to be largely automatic, beyond the control of will, instructions, etc.

This finding, suggesting the existence of an automatic memory mechanism subsequently paved the way for a series of new experiments where changes in basic stimulus features (frequency, intensity, and duration) and the elicitation of the MMN were addressed in more detail. It was established that the MMN is elicited by both increments and decrements in basic stimulus features. The MMN, however, is not elicited when a stimulus sequence begins or, similarly, when stimuli are presented with very long interstimulus intervals (ISIs). Thus, it was concluded that no stimulus per se is an adequate stimulus for the MMN generator mechanism, as the system responds to the difference between the consecutive stimuli. This response pattern is clearly separable from the behavior of N1 response; the N1 amplitude is largest in response to the first stimulus of a series, strongly attenuating thereafter and showing only partial recovery to a subsequent different stimulus.

The observations that even decrements in stimulus intensity and, especially, in duration elicited the MMN need to be stressed because they provide strong evidence that the MMN is not a reflection of "fresh" neural activity in tonotopically and/or ampliotopically organized cortex (i.e., whereby the responses elicited by deviant stimuli would be generated in a location spatially separate from that of responses elicited by standard stimuli).

In the past decade, there have witnessed a resurgence in the electrophysiological literature of attempts to understand how the brain processes the speech signal (Boyd et al., 1996; Buchwald et al., 1994; Dehaene-Lambertz et al., 1994; Dehaene-Lambertz, 1997, Kraus et al., 1993, 1995; Maiste et al., 1995; Molfese, 1987; Näätänen et al., 1997; Phillips et al., 1995; River-Gaxiola et al., 2000;

Sharma et al., 1993; Woods, 1986). One of the most used and well known paradigms in electrosphysiological research is the so-called oddball paradigm, in which typically two stimuli are presented, in random order. One of the stimuli occurs less frequently than the other and the subject is required to discriminate the infrequent stimulus (deviant, target or oddball) from the frequent one (standard). Two main types of ERPs have been described in the literature as a response to the detection of the deviant: P300 (Buchwald et al., 1994; Dehaene-Lambertz et al., 1994; Maiste et al., 1995) and the MMN (Aaltonen et al., 1993; Kraus et al., 1993, 1995; Maiste et al., 1995; Näätä nen R., et al., 1993, 1997; Sams M., et al., 1985; Sharma A., et al., 1993). If the subject is required to respond overtly --- for example, by pressing a button – each time he/she detects the deviant, a positive wave peaking approximately 300 ms after deviant onset is elicited. This wave is called P300 and it is largest over electrode sites in normal adults. Such positivity is thought to reflect voluntary focused attention (context updating, response selection). However, if the subject is not required to respond overtly, and one subtracts the ERPs obtained in response to the standard, from the ERPs obtained for the deviant, a so-called mis-match negativity (MMN) may be observed, usually peaking between 100 and 300 ms after stimulus onset depending on the characteristics of the difference between standard and deviant stimuli. This component is thought to reflect a pre-attentional detection of deviation, a mismatch between the deviant and the memory trace formed by the standard.

Accurate pitch perception is a prerequisite for adequate music and speech processing for instance, in the case of melodic, harmonic, and prosodic processing. Neural dynamics underlying pitch processing have recently been extensively investigated by recording the mismatch negativity (MMN) component of the event-related potential (ERP). The MMN is elicited when the neural code of the incoming stimulus with ‘deviant’ parameter(s) is discrepant with the cortical representation of the repetitive ‘standard’ stimulus (Näätänen, 1992; 1997). The MMN is elicited by any discriminable sound change, for instance, in sinusoidal tones (Näätänen, 1992), phonemes (Aaltonen et al., 1994; Alho et al., 1998; Kraus et al., 1996; Näätänen et al., 1997; Tervaniemi et al., 1999), chords (Alho et al., 1996; Koelsch et al., 1999; Tervaniemi et al., 1999), and also abstract sound features (Paavilainen et al., 1998).

The MMN latency and amplitude reflect the magnitude of the physical difference between the deviant and standard stimulus (Sams et al., 1985). Furthermore, the MMN amplitude and latency reflect perceptual accuracy, as determined by musicality tests (Tervaniemi et al., 1999). Nd by hit rates and the reaction times in a behavioral experiment employing the same stimuli as in the MMN recordings, (Aaltonen et al., 1994; Kraus et al., 1996; Tiitinen et al., 1994). This correlation between the MMN parameters and behavioral responses imply that pre-attentive neural functions determine the accuracy of the subsequent attentive processes (Novak et al., 1990).

However, there exists a discriminative auditory process that does not require the subject's attention. When a sound violates the regularities of the preceding auditory sequence, it elicits an event-related brain potential (ERP) component termed the mismatch negativity (MMN) even when the subject is engaged in some task completely unrelated to the auditory stimulation. The MMN component appears as a frontocentrally negative wave usually peaking between 100 and 300 ms from the onset of stimulus deviation. A major part of the prototypical MMN paradigm, deviant sounds infrequently replacing a repetitive auditory stimulus (standard) elicit an MMN. MMN elicitation in such situations indicates that the deviant sound was pre-attentively discriminated from the standard stimulus represented by its memory trace. MMN is elicited by changes in synthesized as well as naturally- produced speech sounds (Winkler et al., 1999).

A neurophysiological paradigm well suited to examine pre-attentive and automatic central auditory processing is the mismatch negativity (MMN). This is a negative component of the event-related brain potential (ERP), elicited when a detectable change occurs in repetitive homogeneous auditory stimuli (Näätänen, 1992). The most commonly described MMN occurs at 100-300 ms post-stimulus onset although other studies have found later MMNs between 300 and 600 ms (Kraus et al., 1996). The MMN is elicited by any change in frequency, intensity or duration of tone stimuli, as well as by changes in complex stimuli such as phonetic stimuli (Näätänen, 1992). It is assumed to arise as a result of a mechanism that compares each current auditory input with a trace of recent auditory input stored in the auditory memory. The MMN usually reaches its amplitude maximum over the fronto-central scalp (Näätänen, 1992).

MMN is a negative component of the event-related potential (ERP) which is elicited by any discriminable change (Näätänen and Tiitinen, 1998) in a repetitive background of auditory stimulation (Winkler et al., 1996). When deviations in regular auditory stimulation occur, the mismatch negativity (MMN) component of the event-related brain potential (ERP) is elicited pre-attentively Näätänen et al., 1987, Schroger, 1998, Tiitinen et al., 1994). In several hundred studies, the MMN technique has been used as an indicator of an automatic change detection system of the human brain in the processing of auditory information (Näätänen, 1992). There is a strong belief that MMN is elicited if a memory-comparison process detects a discrepancy between the neural representation of the regularity inherent in recent stimulation and the representation of the current stimulus. The traditionally MMN technique yields a combination of memory-comparison-based and memory-comparison-unrelated deviance-related effects (Jacobsen and Schroger, 2001).

Currently, MMN provides the only objective measurement of sound-discrimination accuracy (Näätänen, 2001). There is a close relationship between the MMN amplitude and the discrimination accuracy (Näätänen et al., 1993). Moreover, MMN can also be used to index the training-related improvement in discrimination ability (Näätänen et al., 1993). Because of these properties, MMN has been successfully used in dyslexia research (Baldeweg et al., 1999), for example, for evaluating the effectiveness of different training and rehabilitation programs. Other fields of the potential application of MMN include early language development (Kraus and Cheour. 2000), cochlear implants (Pinton et al., 2000), schizophrenia (Javitt et al., 2000; Michie,2001; Kirino and Inoue 1999), dyslexia (Kujala and Näätä nen, 2001; Schulte-Körne et al., 2001; Balderweg et al., 1999), epilepsy (Liasis et al., 2001), alcoholism (Ahvenine et al., 2000; Grau, 2001), aging (Jääskeläinen et al., 1999), Alzheimer's disease (Pekkonen et al., 1994), and coma monitoring (Morlet et al., 2000) and outcome prediction (Fischer et al., 2000).

CHAPTER II
OBJECTIVE

In this chapter, the main hypotheses that are investigated in this study are outlined. In addition to the experimental work reported in the Results Chapter, the importance of the development of a Mismatch negativity (MMN) using a passive oddball paradigm for the neurophysiological study of preattentive processing in both native and nonnative speakers of Thai is emphasized. The details of this model are presented in the Materials and Methods Chapter.

When engaged in a conversation, listeners tune in to the relevant stream of speech and filter out irrelevant speech input that may be present in the same environment. Nonetheless, attention might be involuntarily diverted to meaningful items coming from an ignored stream, like in the well-known own-name effect (Moray, 1959). This brings up the question of to what extent speech is processed in the ignored streams. A fundamental, yet controversial issue is whether the human brain contains neural circuits that are uniquely engaged in the early, pre-attentive stage of speech processing. While it seems indisputable that language is subserved by left-hemisphere (LH) and right-hemisphere (RH) are lateralized for speech, language, or something else. Hypotheses proposed to account for functional hemispheric asymmetries can generally be classified as either *cue dependent* i.e., basic neural mechanism underlie processing of complex auditory stimuli regardless of linguistic relevance (Ivry and Lebby, 1993), or *task dependent*, i.e., specialized neural mechanisms exist that are activated only by speech (Imaizumi et al., 1998).

The present study thus proposes a task-dependent hypotheses in which the phonological units of segmental (e.g., consonant and vowel) and suprasegmental (e.g., prosodic) are processed in discrete asymmetry of the human brain. Although the left hemisphere (LH) is selectively employed for processing linguistic information irrespectively of acoustic cues or subtype of phonological unit, the right hemisphere (RH) is employed for prosody-specific cues (Imaizumi et al., 1998). The propose of the present study is, thus, to use both an auditory MMN component of event-related potential (ERP) recording and the low resolution electromagnetic tomography (LORETA) techniques to measure the degree of cortical activation and to localize the brain area contributing to the scalp recorded auditory MMN component, respectively, during the passive oddball paradigm. Thus, the specific objectives of the work reported here in this study are outlined below:

Hypothesis 1: Mismatch negativity (MMN), a unique indicator of automatic cerebral processing of acoustic stimuli, can be used to provide an index of experience-dependent and long-term memory traces for category and duration of vowel in Thai words.

The human voice recognition and discrimination, is amongst the most important functions of the human auditory system, has been recently measured from the electrophysiological activity of the perceiver's brain (Titova and Näätänen, 2001). This was investigated by using an objective measure of pre-attentive sound discriminability, called the mismatch negativity (MMN), a component of the auditory event-related (ERP) (Näätänen and Alho, 1995). Mismatch negativity (MMN) can be used to investigate the neural processing of speech and language (Näätänen, 2001; Näätänen and Winkler, 1999; Pulvermüller et al., 2001; Shtyrov et al., 2000; Näätä nen, 1999; Näätänen et al., 1997; Alho et al., 1998; Shtyrov et al., 1998) because it is considered to be a unique indicator of automatic cerebral processing of acoustic stimuli (Shtyrov and Pulvermüller, 2002). MMN, with its major source of activity in the supratemporal auditory cortex, is a brain response elicited in an oddball paradigm where a sequence of repetitive, 'standard', stimuli is interspersed with occasional 'deviant' stimuli that differ from the standard in one or several acoustical or temporal features (Alho, 1995; Picton et al., 2000; Cowan et al., 1993). MMN is thus primarily a response to an acoustic change and an index of sensory memory. Importantly, the MMN can be elicited in the absence of the subject's attention (Näätänen, 1995).

The MMN component appears as a frontocentrally negative wave usually peaking between 100 and 300 ms from the onset of stimulus deviation. The MMN(m) in response to speech sounds can occur as a response to both within- and across-category change of phonemes. The former is the change detection of physical feature(s) within the same phoneme boundary, such as changes in pitch, duration, and intensity. The latter is the change detection between independent phonemes (Kasai et al., 2001). For example, Thai language can be divided into front (i.e. /i/, /e/, /ɛ/, /æ/) and back (i.e. /M/, /o/, /O/, /a/) vowels or rounded (i.e. /u/, /o/, /O/) and unrounded (i.e., /i/, /e/, /ɛ/, /æ/, /a/) vowels. Consequently, the change detection between any combinations of these (i.e., /O/ *vs.* /i/) is considered across-category change detection.

The MMN is traditionally thought to reflect the acoustic difference between two sounds, so that the larger the acoustic difference, the faster the behavioral responses of the listeners' and the earlier and larger the MMN. However, if the deviant vowel is near the category boundary, then the familiarity of the stimuli modulates the reaction time (Savela et al., 2003). Moreover, the reaction time results on the speech sound categorization and discrimination demonstrate that vowels were categorically coded (Pisoni, 1973; Polka, 1995). This indicates that across-category discrimination performance is more accurate than within-category discrimination.

Because the vowel categorization is based on several simultaneous encoding mechanisms, numerous attempts have been made to establish the underlying cortical functions. MEG data of the auditory evoked field show that the different vowels elicited distinct spatio-temporal pattern of cortical activity and were represented by different cortical mappings, as can be expected by their different spectral structures (Mäkelä et al., 2003; Obleser et al., 2002). Vowels with mutually exclusive place of articulation features elicited separate centers of activation in the vicinity of the auditory cortex (Obleser et al., 2002). For example, the acoustically most dissimilar vowels /a/ and /i/ showed more distant location than the more similar vowels /e/ and /i/ (Obleser et al., 2003). The distance between sources representing /u/ and /i/ was larger than the distance between /a/ and /ɛ/ (Diesch et al., 1996). Thus, the asymmetries in vowel perception evident in human are likely to reflect general auditory mechanisms and can be considered indicative for a cortical map related to spectral dissimilarities of speech input (Obleser et al., 2003; Polka and Bohn, 2003).

Furthermore, the left and right hemispheres of the brain are not symmetrically involved in language processing. The human brain processes speech-relevant acoustic information in both the left and right temporal lobes especially in the superior temporal gyrus which contains neural networks specialized for language processing (Jäncke et al., 2002). The present study was thus hypothesized that the two hemispheres may have two time courses for activation: (1) both hemispheres are active in parallel with the left hemisphere dominant for different phonemes throughout and (2) hemispheric dominance may alternate to play a specific role at different stages of linguistic processing. Both hemispheres may symmetrically activate in the preattentive across-category change perception of vowels while the within-category change of a vowel was predominantly activated in the right hemisphere. Additionally, the dominance of the left auditory cortex in the preattentive speech processing might occur only at the level of perception of vowel across-category change.

Hypothesis 2: The preattentive processing can provide an index of hemispheric specialization of vowel duration and tone changes in Thai words.

In an earlier studies (Gandour et al., 1998, 2000; Hsieh et al., 2001), Chinese and English listeners did not show the same left-hemisphere (LH) lateralization as Thai listeners when making perceptual judgments of Thai tones. In addition, Chinese and English listeners were asked to make perceptual judgments of Chinese tones, consonants, and vowels. Chinese listeners showed left-hemisphere (LH) lateralization for both suprasegmental and segmental phonological units (Hsieh et al., 2001). These earlier studies suggest that functional circuits engage in early, pre-attentive speech perception of either suprasegmental or segmental units in tone languages.

Previous ERP studies at a phonetic level demonstrated that the MMN was enhanced in Finnish subjects by their first-language (Finnish) phoneme prototype rather than a non-prototype (Estonian) (Näätänen et al., 1997) and that the MMN for a vowel contrast in Finnish was not generated in native Hungarian speakers with no knowledge of Finnish (Winkler et al., 1999), implying that the MMN reflects language-specific memory traces formed by early and extensive exposure to a first language. However, language-specific word-related MMN/MMF components at acoustic and phonetic levels remain to be investigated in future studies. The differences between these studies provide the impetus for future investigations of duration processing and temporal integration differences across language groups.

This present study may support previous results showing that MMN was elicited by both stimulus duration increments and decrements (Jaramillo et al., 1999; Joutsiniemi et al., 1998; Näätänen et al., 1989). MMN generally increased as a function of the degree of discrepancy from the standard-stimulus duration. Nonetheless, the fact that the MMN was still elicited by such an extremely brief sound suggests that the MMN generator is indeed sensitive to even very small amounts of stimulus energy. Consequently, the finding that the MMN amplitude increases as a function of discrepancy in duration and tone change from the standard suggests that the MMN is generated by a process which is a neural code of the degree of stimulus change and accord with results from similar paradigms using other features (Näätänen et al., 1989).

Hypothesis 3: The preattentive processing can provide an index of hemispheric specialization for cluster and noncluster consonants of Thai words.

In the present study, the MMN was recorded from both NS and NonS in responses to the synthesized stimuli /k/ and /kl/ as phonemes and as in CV syllables. As it is well established that the MMN amplitude indexes the change detection accuracy (Näätänen, 2001), the larger MMN amplitude to the phonetic change in the native speakers suggests more accurate phonetic change detection in native speaker than in the second-language users. It is proposed that the tuned processing of phonetic sound in the native speaker and the lack of it in the second-language users is caused by the different roles of sound frequency and duration in the subjects' native language. Hence the fine tuning in the processing of the phonetic sound may be inhibited at the pre-attentive level.

The present study employed the mismatch negativity (MMN) approach which enables us to distinguish brain activation patterns attributable directly to the early, pre-attentive sound discrimination processing of the *suprasegmental (prosodic) phonological* unit from those attentive sound discrimination processing and independent complex-sound processing. The hypothesis will support the domain-general issue that neurobiological mechanisms are specialized for complex aspects of auditory processing, and, furthermore, that they are sufficient to account for the lateralization of linguistic-prosodic functions to the right hemisphere (RH) in the early, pre-attentive pitch perception.

CHAPTER III
MATERIAL AND METHODS

This chapter describes briefly the electrophysiological and brain electromagnetic tomography procedures used to prepare the passive oddball paradigm for the preattentive processing study.

SECTION 1: SUBJECTS AND PROTOCOL OF THE EXPERIMENTS

3.1. Subjects

In each experiment, twenty-two (except Experiment 1) healthy right-handed (Handedness assessed according to Oldfield (1970) adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 18-35 and eleven NonS, aged 23-35. All participants gave their written informed consent before participation in the study. The mean (±s.d.) age was 24.35 (± 4.95) years.

3.2. Handedness

The handedness of the subjects will be assessed with the Edinburgh Handedness Inventory. The degree of the right handedness of the subjects will be assessed based upon ten items; writing; drawing; throwing; scissors; toothbrush; knife (without fork); spoon; broom; striking a match; and open box lid. The subject will be instructed to make a “ + “ on which hand he/she would prefer to use for each action. The subject will be instructed to mark a “ ++ “ when the preference is so strong that he/she never used the other hand unassisted. If, in any case, the subject does not have any preference, he/she will be instructed to mark a “ + “ for both hands. The numbers of “ + “ marked for each hand will be totaled. Then, a handedness index will be calculated to be the difference of the numbers of “ + “’s between the right and left hands divided by the total number of “ + “’s for both hands. A handedness index of 1.0 indicated completely right handed, -1.0 corresponded to completely left handed, and 0 suggested ambidextrous. The subject is also asked which foot is preferred for kicking, which eye is preferred when only using one eye, and whether both parents are right handed.

3.3. Ethical Considerations

All participants gave their written informed consent to participate in the experiments and were paid for their participation. The experiments were performed in accordance with the Helsinki Declaration. Ethical permission for the experiments was issued by the Committee on Human Rights Related to Human experimentation (Mahidol University, Thailand).

3.4. Language

It is a common assumption in research on speech perception that a listener's internal representation of the speech signal is organized, at least in part, in terms of phonetic and/or phonological features employed in linguistic analysis (Gandour, 1979). The pitch of complex sounds corresponds to the fundamental frequency (Fo) that is derived from the combination of two or more separate spectral components (Braum, 2000). The process of speech perception has been summarized by Studdert-Kennedy (1975) summarized the process of speech perception that perception entails the analysis of the acoustic syllable, by means of its acoustic features, into the abstract perceptual structure of features and phonemes that characterize the morpheme.

In general, the understanding of speech sounds, to the normal listener, depend upon three important factors: first, the hearing acuity at various test-frequencies; second, the neuro-sensory receptor for hearing discrimination; and third, individual speech and language development. When the neuro-sensory function is damaged, the speech discrimination ability is decreased. Hearing-impaired listeners always confuse consonant and vowel phonemes, which are commonly found in reports.

Almost all languages have two kinds of phoneme: consonant and vowel. In tone languages, phonemic tones free of a linguistic context are better identified when the listener has access to the speaker's tone space. The listener may need information to identify words minimally differentiated by phonological relevant tones. Relative, rather than absolute, values of the fundamental frequency (Fo) of the voice are normally found to provide the major acoustic cues for identification of phonemic tones even in the presence of other phonetic features (e.g. creaky voice, voice breaking, and amplitude variations). Certain tones of a given system may be quite identifiable in isolated citation forms of syllables, while others may have to be embedded in a linguistic context for easy recognition. Even the latter tones, however, may enjoy high identifiability in isolation if the listener is free to adapt his perception to the tone system of a given speaker. The perceptual assessment of the tones of a language may well be analogous to that a vowel space (Abramson, 1976).

Tone languages (e.g., Thai, Chinese, and Vietnamese) are ideally suited for providing a condition in which pitch is phonologically significant, thus, enabling experimental comparison of pitch in nonlinguistic and linguistic contexts. In tone languages, pitch variations signal lexical contrast (for example, Thai: /ma:/ 'come' vs. /mă:/ 'dog'). Since variations in pitch patterns at the syllable level are linguistically significant, experimental manipulations with tone languages offer a unique opportunity to elucidate the functional neuroanatomy of speech prosody (Gandour et al., 2000).

In the perceptual dimensions of Thai tones, there are four dimensions which underlined the person's dissimilarity rating between pairs of pitch patterns superimposed on a synthetic speech-like syllable. Four dimensions are interpreted as Average Pitch, Length, Direction, and Slope. Of the four dimensions, Average Pitch is the most important factor that ones incorporate into their perceptual judgments. The perceptual of tone also deals with how a listener transforms, organizes and structures the pitch information arising from the speech signal (Gandour, 1979).

In the Thai language, tone is another significant phoneme in the language system which listeners have to learn to discriminate. Thus, the addition of tonal confusion may be found in hearing-impaired Thai listeners, since its acoustic properties and closely related to frequency, which will be perceived by the ear. So, it is interesting to know how much of an important role hearing acuity will play in the perceptual ability of Thai tones and how hearing-impaired Thai listeners perceive the tones (Nimitbunnasarn, 1984). Many studies have revealed that Thai with normal listening ability can easily identify isolated monosyllabic words with phonemic tones and with a high degree of accuracy (Abramson, 1962, 1972, 1975, 1976). Thai language, the national language of Thailand, exhibits five contrastive lexical tones (Gandour, 1975), traditionally labeled mid (   ), low (   ), falling (   ), high (   ) and rising (   ): for example, /k^{h}aà/ ‘galangal’, /k^{h}aâ/ ‘kill’, /k^{h}aá/ ‘trade’, /k^{h}aă/ ‘leg’. The mid tone can be described phonetically as midlevel with a final drop, low tone as low-falling, falling tone as high-falling, high tone as high-rising, and rising tone as low-rising.

SECTION 2: ELECTROPHYSIOLOGICAL PREPARATIONS

3.5. General Electrophysiological Procedures

Subjects are instructed to sit relaxing in comfortable reclining chair in an electrically and acoustically dampened room at the Clinical and Research Electroneurophysiology Laboratory, Neuro-Behavioural Biology Center, Mahidol University, Thailand. They are told that they would participate in a speech perception experiment and that the experimenter would be recording their brain electrical activity. They are given written instructions and provided with a grid for their judgements and a pen. They silently read the instructions and at the end the experimenter verifies that everything is clear. Their histories are taken, including age, educational level, handedness, occupation, current medications, medical history (which included past illness, surgical history, head trauma or accident) and history of alcohol consumption or smoking. If there is any significant history of neurological problems, psychiatric problems or head trauma, that participant is excluded. Before the EEG/ERP data is obtained, memory and language capacities are assessed with the Digit Span subtest of the Weschler Adult Intelligence Scale and the Token Test, respectively.

For the Mismatch Negativity (MMN) study, all subjects are instructed to ignore the tone stimuli by watching a silent, subtitled video of their choice (ignore condition). They are asked to avoid body and eye movements and to keep alert. A practice session is given in two phases. In the first phase, both acoustical phonetic stimuli sequence of monosyllabic Thai words are presented one by one at the rate of stimulation the tones occurred in the alternative sequence. Subjects are instructed to listen for all tones sequence and tell what they heard. When it is clear that the subjects could hear all tones, the second phase is conducted. In the second phase of the practice session, an alternating sequence of tones similar to that to be used during the recording session is presented. Subjects are now instructed to ignore the tone sequences. When subjects understand what to do and could perform the task, recording proceeded. Before the recording session, the task is explained and a practice block of 50 tones (50 deviants) is presented to the participant to ensure a good level performance. In order to avoid alpha rhythm synchronization during the recording session, subjects are instructed to remain with their eyes open while watching a silent, subtitled video of their choice and are instructed to avoid eye movement and blinking. The total experimental session is 3 h. including approximately 0.20 h. for electrode placement. During the experimental session, subjects take a rest breaks (one 15-min break occurring halfway through the recording session and shorter 5-min breaks as needed). Subjects are tested in all experimental conditions on the same day.

3.6. Stimuli and procedure

Three consonant-vowel syllables of Thai were prepared: back rounded articulation /pO/, front unrounded: /pi̲/, and back rounded: /po̲/. All stimuli were spoken by native female Thai speaker and digitally edited to have an equal peak energy level in dB SPL using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation) with 500 ms duration. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB (determined using a Brüel & Kjaer 2230 sound level meter). All stimuli were identical at their suprasegmental (i.e., tone) unit, which was always “mid” tone, thus eliminating any effect due to differences in frequency of occurrence of tones. The /pi/ (10%) (across-category vowel change) and /po/ (10%) (within-category change vowel) deviants were presented among the /pO/ standard (90%) in random order (except that each deviant stimulus was preceded by at least one standard stimulus). The inter-stimulus interval (ISI) was 1.25 second (offset-onset).

Additionally, eight different stimuli were synthetically generated: (1) /k^haam/ - long vowel, level tone; (2) /k^haam/ - long, falling, (3) /k^haam/ - long, rising, (4) /k^ham/ - short, level, (5) /k^ham/ - short, falling; (6) /k^ham/ - short, rising; (7) /kaang/ - noncluster initial consonant; (8) /klaang/ - cluster initial consonant. The sound-duration difference between stimuli are vary, but with the same intensity used in each stimulus. Five native speakers of Thai listened to the synthesized words and evaluated them all as natural sounding.

3.7. Electroencephalographic recording

For EEG/ERP recording, the standard 20 locations of the 10-20 system (Jaspers, 1958), EEG is recorded via an electrocap (Electrocap International) from 20 active electrodes (Fp1, Fp2, F7, F3, Fz, F4, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz, O2) positioned according to the 10-20 International System of Electrode Placement, plus Oz and Ground are applied, pre-mounted in an elastic Electro-Cap. Reference electrodes are manually applied to left and right mandibles, where the Fp1 and Fp2 electrodes are used for ocular artifact detection. Horizontal eye movements are monitored with electrodes at the left and right outer canthi and vertical eye movements are monitored at Fp1 and Fp2. EEG is amplified with a gain of 30,000 and filtered with a bandpass of 0.1-30 Hz. EEGs are acquired as continuous signals and are subsequently segmented into epochs of 1s (a 100 ms pre-stimulus baseline and a 900 ms post-stimulus epoch).

The electrodes are then filled-up with conductive jelly, the impedance of electrode skin contact for all channels is tested and maintained below 5 k and all have approximately the same range of impedance values. The electrical activity is amplified with the built-in DC-coupled amplifiers in the Brain Atlas III (Bio-logic Systems Corp., Mundelein, Illinois, U.S.A.) and filtered with a bandpass of 10-30 Hz (notch-filter at 50 Hz), for all referential EEG and ERP recordings. Sampling rate for all signals is 128 Hz. A short run each of EEG recording with eyes-open and eyes-closed is performed to rule out any types of abnormal EEG that could affect the results of the study, such as slowing, sharp waves, etc. This is performed for at least 30 seconds of artifact-free EEG recording in each condition. Subjects' data are transferred from the scalp electrodes to the Brain Atlas III's amplifiers.

For ERPs recording, a total of 500 sweeps to standard tones and 125 sweeps to deviant tones are obtained for off-line averaging of the signal. Epochs with deflections larger than 100 μV at any of the Fp1 or Fp2 channels are rejected from averaging because they are probably contaminated by extra-cerebral artifacts such as eye movements, blinks, or muscle activity. Responses with high-voltages, such as muscle artifact, eye-blinks or eye-movement, are automatically rejected by the averaging computer using a 'voltage gates' system. Artifact rejection is set to omit activity-exceeding ± 100 μV for all electrodes to exclude blinks and movement artifacts. Each does at least two trials, with a third being necessary in a few cases, due to poor problems of poor waveform replication, etc., until a reliable result is obtained. During the recording, the computer continuously updates the averaged response for operator viewing. Since the computer also controls stimulus delivery, it would continue stimulus presentation until 125 artifact-free responses to deviant stimuli have been obtained for each trial. These 2 replicating trials are combined together afterwards in the Bank Mathematics part of the software program and normalized to yield a Grand Average waveform for data analysis. The ERPs were separately averaged for the standards and deviants (omitting epochs with EEG or EOG artifacts exceeding ± 100 μV) and digitally filtered (bandpass 0.1-30 Hz). In order to get the maximal Mismatch

Negativity (MMN) amplitudes at the frontal (Fz) electrode site, the ignore condition ERPs are arithmetically re-referenced to the average of the mandibles as the Mismatch Negativity (MMN) inverts its polarity at electrodes below the level of the Sylvian fissure (Näätänen, 1995 and Ritter et al., 1992).

The ERPs are averaged for 1024 ms, including 100 ms of pre-stimulus baseline.

Epochs time-locked to the stimuli start 100 ms before (pre-stimulus baseline) and ended 900 ms after stimulus onset. They are averaged on-line separately for standard and deviant stimuli. The averaged EEG responses are digitally filtered. For data illustration, the pass band of 10-30 Hz is used. Sampling rate for all signals is 128 Hz, following which the ERP data is converted from the time domain to the frequency domain by Fast Fourier Transformation (FFT), producing 64 two-dimensional color maps, representing frequencies between 0 – 31.5 Hz, spaced 0.5 Hz apart. The display points are computerized using a 'nearest four point' linear interpolation algorithm, with 'Cold' (blue) colors representing low amplitudes and 'Warm' (red) colors representing high amplitudes. Values are expressed in microvolts.

3.8. EEG data processing

The recordings were filtered and carefully inspected for eye movement and muscle artifacts. ERPs were obtained by averaging epoch, which started 100 ms before the stimulus onset and ended 900 ms thereafter; the -100-0 ms interval was used as a baseline. Epochs with voltage variation exceeding ±100 μV at any EEG channel were rejected from further analysis. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. For each subject, the averaged MMN responses contained at least 125 accepted deviant trials. All responses were recalculated offline against average reference for further analysis.

3.9. Instrumentation

Bio-logic Systems 'Brain Atlas III' Topographic Brain Mapping

Computer Amplifiers - 20 built-in DC-coupled amplifiers, including A/D Converter

Electrodes - Metal (pure tin) 'recessed' type electrodes in elastic cap ('Electro-Cap')

Electrolyte - High-viscosity EEG gel (Siriraj Hospital Pharmacy, Bangkoknoi, Bangkok)

Electrode Placement - International 10-20 System of Electrode Placement (Jasper, 1958)

Electrode Impedance - 5 K (5,000 Ohms) or less

Number of electrodes - 19 standard 10-20 positions plus Oz, for a total of 20 electrodes; Fpz position interpolated; Ground on the high forehead

Montage - Referential to Linked Mandibles; re-referenceable to the Common Average

Calibration - BA-III Calibration procedure automatically tests for voltage output and DC offset; subsequent independent channel correction under software control

Amplification - Gain set at 30,000

Filter band-pass - 1-30 Hz, 50 Hz 'notch' filters 'ON'

3.9. Apparatus

All of the natural acoustical phonetic stimuli are digitally edited to have equal peak energy level in dB SPL with the remaining data within each of the stimuli scaled according by the CoolEdit 2002 program (Syntrillium Software Corporation) and WinCecil program (Summer Institute of Linguistics). The sound pressure level of all natural acoustical phonetic stimuli are then measured at the output of the earphones (Telephonic TDH-39-P) in dBA using a Brüel & Kjaer 2230 sound level meter. All natural acoustical phonetic stimuli are delivered from a PC using the SuperLab software (Cedrus Corporation, San Pedro, USA). Intensity levels are 85 dB SPL. All natural acoustical phonetic stimuli are then presented binaurally via form insert headphones (Panasonic RP-HT 222) at a comfortable listening intensity level of approximately 85 dBA SPL. The experiment is run in an electrically and acoustically dampened room at the Clinical and Research Electroneurophysiology Laboratory, Neuro-Behavioural Biology Center, Mahidol University, Thailand. EEG is recorded using mental (pure tin) 'recessed' type electrodes in elastic cap ('Electro-Cap'), 20 built-in DC-coupled amplifiers, including A/D converters, a 20 –channel Brain Atlas III Topographic Brain Mapping unit manufactured by Bio-logic Systems Corp, Mundelein, Illinois, USA. Off-line signal processing is carried out in the Bank Mathematics part of the software program and normalized to yield a Grand Average waveform for data analysis.

SECTION 3: SPATIAL DATA ACQUISITIONS

3.11. Mismatch Negativity (MMN) Peak-Picking Procedure

In order to visualize and to measure the MMN (deviant-tone ERP-minus-standard-tone ERP difference), after the recording, differences are calculated by subtracting the ERP elicited by the standard tones from that elicited by the corresponding deviant tones of the same stimulus class. The MMN is quantified by first determining the MMN peak latency from the frontal (Fz) grand-average difference waves separately for each deviant. The latency windows for picking up the MMN peaks are predefined on the basis of the across-participants peak latency distribution, determined by visual inspection.

The MMN component is defined as the most prominent negative peak within the time windows between 100 and 300 ms. Latency and amplitude figures for waveforms are picked at their point of maximal deflection, as seen at their electrode site of maximal voltage distribution of frontal (Fz) electrode site. Peak-picking of the prominent peak (MMN) is accomplished by means of moving an 'enhanced point' cursor through the waveforms displayed on the computer screen, while simultaneously paying attention to the resultant changes in the topographic maps. The mean MMN amplitudes at the frontal (Fz) electrode site are calculated as a mean voltage of the 40 ms intervals (so the peak plus minus 20 ms), centered at the corresponding peak latencies of the left and right frontal electrodes in the grand-averaged waveforms, separately for each stimulus type. The amplitudes are determined by using the 100 ms pre-stimulus baseline. When the participants are watching a silent, subtitled video, MMN to spatial acoustic changes was observed as a significant difference between ERPs to the deviant tones and those to the standard tones. It is at its maximum at the frontal (Fz) electrode site consisting of a negative deflection (note that analyses are based on averaged 40-ms blocks of sample points). MMN amplitudes are measured as the mean amplitude over the 100-300 ms period after the stimulus-onset from the deviant-tone ERP-minus-standard-tone ERP differences.

3.12. Spatial analysis

The average MMN latency was defined as a moment of the global field power with an epoch of 40-ms time window related stable scalp-potential topography (Pascual-Marqui, 1994). In the next step, low-resolution electromagnetic tomography (LORETA) is applied to estimate the current source density distribution in the brain, which contributes to the electrical scalp field (Pascual-Marqui et al., 1994, 1995). Maps are computed with the Low Resolution Electromagnetic Tomography (LORETA). This study selects two radically oriented point sources (dipoles) in the brain and computes the 21 channel forward solution electric potential map using a 3-shell unit radius spherical head model (Ary et al., 1981). The forward solution maps are then used as input for the LORETA computation in order to test the location precision and the ability of the method to separate the two known dipole locations. Scalp potentials re-referenced to the average reference, excluding the EOG electrodes, are interpolated for mapping using the surface spline method. The CSD maps are computed with the spherical spline interpolated data. The maps are computed at a single time point where the component in question is largest in the grand mean waveforms of each stimulus type and condition separately. LORETA computes the smoothest of all possible source configurations throughout the brain volume by minimizing the total squared Laplacian of source strengths.

The individual momentary potential measures from 21 electrodes at the MMN latency were analyzed with LORETA to determine the MMN source loci Pascual-Marqui, 1994). These latencies were between 100-140 ms for across- and within-category changes. LORETA calculated the current source density distribution in the brain, which contributed to the electrical scalp field, at each of 2395 voxels in the gray matter and the hippocampus of a reference brain (MNI 305, Brain Imaging Centre, Montreal Neurological Institute) based on the linear weighted sum of the scalp electric potentials (Pascual-Marqui, 1994). LORETA chooses the smoothest of all possible current density configurations throughout the brain volume by minimizing the total squared Laplacian of source strengths. This procedure only implicates that neighboring voxels should have a maximally similar electrical activity, no other assumptions were made. The applied version of LORETA used a three-shell spherical head model registered to the Talairach space and calculated the three-dimensional localization of the electrical sources contributing to the electrical scalp filed for all subjects and conditions, defining the regions of interest on the basis of local maxima of the LORETA distribution.

Stereotaxic coordinates of the voxels of the local maxima were determined within areas of significant relative change associated with the tasks. The anatomical localization of these local maxima was assessed with reference to the standard Stereotaxic atlas, and validation of this method of localization was obtained by superimposition of the SPM maps on a standard MRI brain provided by the SPM99. Peaks located within superior temporal gyrus was also identified by using published probability maps following a correction for the differences in the coordinate systems between the Talairach and Tournoux atlas and the Stereotaxic space employed by SPM99.

3.13. LORETA and Dipole Computations

Low-resolution Electromagnetic Tomography (LORETA) is the new implementation of LORETA in the Talairach brain. LORETA makes use of the three-shell spherical head model registered to the Talairach human brain atlas (Talairach and Tournoux, 1988), available as a digitized MRI from the Brain Imaging Center, Montreal Neurologic Institute. Registration between spherical and realistic head geometry use EEG electrode coordinates reported by Towle et al. (1993). The solution space is restricted to cortical gray matter and hippocampus, as determined by the corresponding digitized Probability Atlas also available from Brain Imaging Center, Montreal Neurologic Institute. A voxel is labeled as gray matter if it meet the following three conditions: its probability of being gray matter is higher than that of being white matter, its probability of being gray matter is higher than that of being cerebrospinal fluid, and its probability of being gray matter is higher than 33%. Only gray matter voxels at 7-mm spatial resolution are produced under these neuroanatomical constraints. LORETA computations use the exact head model determined from each individual subject's MRI. The final step in any analysis procedure would be to cross-register the individual's anatomical and functional image to the standard Talairach atlas.

3.14. Brodmann Area (BA) and Brain Regions Localization Procedure

Regarding to the Brodmann areas(s) and brain regions localization, the Talairach Daemon (TD) will be taken into consideration. The Talairach Daemon (TD) is a high-speed database server for querying and retrieving data about human brain structure over the Internet (http://ric.uthscsa.edu/td_applet/). The TD server data is searched using x-y-z coordinates resolved to 1x1x1 mm volume elements within a standardized stereotaxic space. An array, indexed by x-y-z coordinates, that spans 170 mm (x), 210 mm (y) and 200 mm (z), provides high-speed access to data. Array dimensions were selected to be approximately 25% larger than those of the Co-planar Stereotaxic Atlas of the Human Brain (Talairach and Tournoux, 1988). Coordinates tracked by the TD server are spatially consistent with the Talairach Atlas. Each array location stores a pointer to a relation record that holds data describing what is present at the corresponding coordinate. Presently, the data in relation records are either Structure Probability Maps (SP Maps) or Talairach Atlas Labels, though others can be easily added. The relation records are implemented as linked lists to names and values for brain structures. The TD server is run on a Sun SPARCstation 20 with 200 Mbytes of memory. Intention is to provide 24-hour access to the data using a variety of client applications, as well as continue to add more brain structure information to the database.

SECTION 4: STATISTICAL EVALUATION

3.15. Mismatch Negativity (MMN) Analyses

The statistical significance of MMN (deviant-minus-standard difference) is tested with one-sample *t*-tests by comparing the mean MMN amplitude at the frontal (Fz) electrode site, where the MMN is most prominent. The MMN is measured using the mean frontal (Fz) amplitude in the 100 - 300 ms interval of the deviant-minus-standard difference curves. This interval includes the grand mean MMN peak latencies in those conditions where MMN is elicited. One-sample *t*- tests are used to verify the presence of the MMN component, by comparing the mean amplitude of the 100 - 300 ms interval against a hypothetical zero, separately in each condition. Statistically evaluation of the data is also performed by one way analysis of variance (ANOVA) in order to determine whether the factors of condition (2: Native and Non-native or Consonant and Vowel or Tone and Pitch) and the different sex groups (2: Male and Female) effected the MMN amplitude and latency at the frontal (Fz) electrode site using the GraphPad Instat program V. 2.02 (GraphPad Software Inc., 1993). All results are expressed as mean ± S.D and all significant

3.16. Voxel-by-Voxel LORETA Images Analyses

For the LORETA analyses, the average LORETA images are constructed across subjects: the brain electric activity during the ERPs amplitude waveforms for each condition and the voxel-by-voxel *t* test differences between conditions. The voxel-by-voxel paired *t* tests are run to assess in which cortical regions the conditions and subject groups differed. The *t* maps are threshold at $p < 0.0001$. Indeed, as pointed out above, reliable differences in the scalp ERP field configuration can unambiguously be interpreted as suggesting that at least partially different neuronal populations are active during the conditions and subject groups. LORETA assesses in which brain regions the conditions and subject groups differed. The Structure-Probability Maps Atlas (Lancaster et al., 1997) is used to determine which brain regions are involved in differences between conditions and subject groups. Brodmann area(s) (BA) and brain regions closet to the observed locations is identified by the Tarairach coordinates are reported. Overall, one sample *P*-values are reported.

CHAPTER IV

RESULTS

This chapter presents results obtained by electrophysiological recording from both native (NS) and nonnative speaker (NonS) of Thai during a passive oddball paradigm. The basic theme is the electric response of preattentive processing to auditory stimulation of acoustic stimuli to both NonS and NS and the investigation of factors influencing these responses. The experiments are designed to test the hypotheses presented in the Objectives Chapter. Studies on enhancing experience-dependent memory traces for category and duration of vowel, consonant and tone changes of lexical tones (e.g., Thai) are described, followed by evidence that electric brain activities response to conditioned stimuli associated with native and non-native linguistic experiences. Lateralization of segmental (e.g., consonant and vowel) and suprasegmental (e.g., tone) phonological units in the brain as indexed by preattentive processing are also shown. Finally, the localization of Brodmann areas (BAs) in the brain are presented.

4.1. Human auditory evoked mismatch negativity amplitudes vary as a function of category and duration of vowels

4.1.1. Experiment 1: Preattentive discrimination of different categories of vowels

4.1.1.1. Subjects

Ten healthy right-handed (Handedness assessed according to Oldfield (1971)) native speakers of Thai (7 females; aged 18-35 years) with normal hearing sensitivity gave their written informed consent before participation in the study (see appendix I). The mean (±s.d.) age was 24.35 (± 4.95) years.

4.1.1.2. Stimuli and procedure

Three consonant-vowel syllables of Thai were prepared: back rounded articulation /pO/, front unrounded: /pi/, and back rounded: /po/. All stimuli were spoken by native female Thai speaker and digitally edited to have an equal peak energy level in dB SPL using the Cool Edit Pro v. 2.0 (Syntrillium Software

Corporation) with 500 ms duration. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB (determined using a Brüel & Kjaer 2230 sound level meter). All stimuli were identical at their suprasegmental (i.e., tone) unit, which was always "mid" tone, thus eliminating any effect due to differences in frequency of occurrence of tones. The /pi/ (10%) (across-category vowel change) and /po/ (10%) (within-category change vowel) deviants were presented among the /pO/ standard (90%) in random order (except that each deviant stimulus was preceded by at least one standard stimulus). The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

4.1.1.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.1.1.4. Results

The grand-averaged ERPs in Fig. 1 show that both across- and within-category vowels changes elicited MMN between 100-140 ms with reference to the standard-stimulus ERPs. The MMN amplitude was statistically significant for both across- and within-category vowel change (Table 1). The MMN amplitude differed significantly between the across- and within-category vowel change conditions being larger in amplitude after an across- than a within-category vowel change (t-test; $t\ (9) = 3.968$; $p < 0.003$).

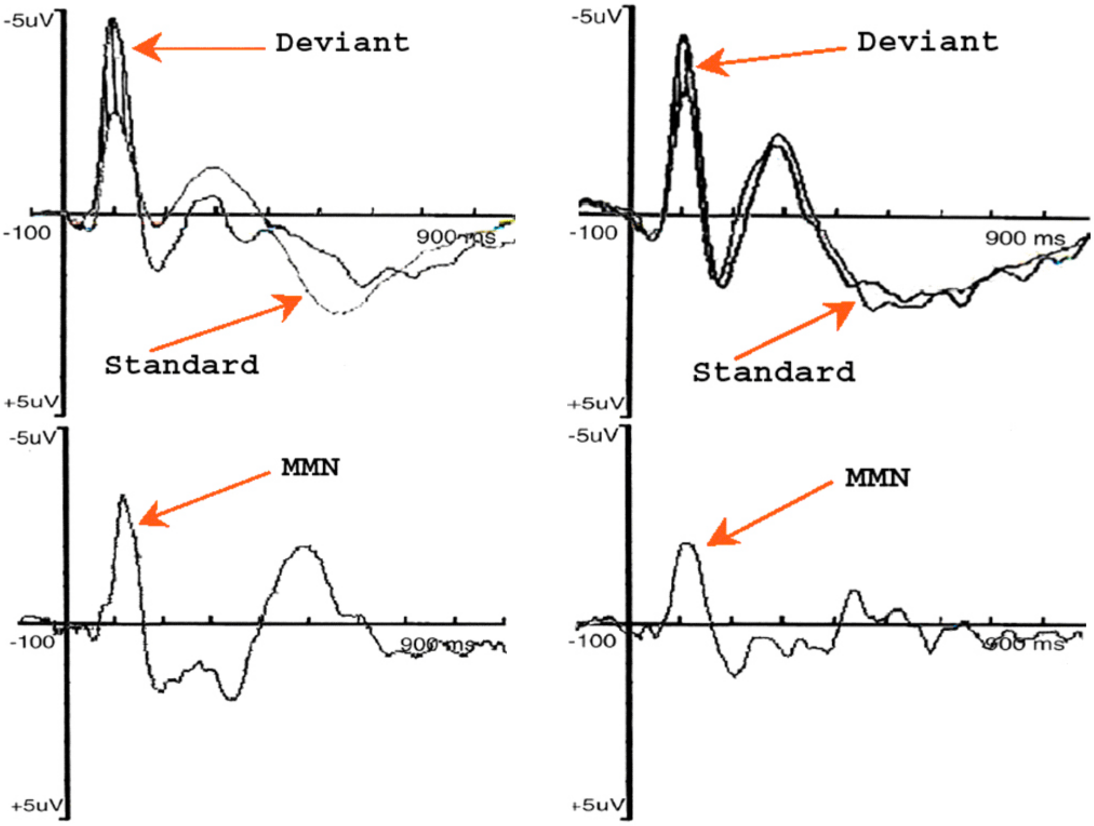

Figure 1: Grand-average event-related potentials elicited by standard and deviant stimuli (top), and deviant-minus-standard difference waves (bottom) at Fz for across-category vowel change:/pO/

Table 1 The MMN mean amplitudes, standard deviations, and *t*-values for the different deviant stimuli used. $P < 0.001$

The deviant stimulus	Mean MMN Amplitude (µV)	Standard deviation (µV)	*t*
Across-category	-3.53	0.90	12.05
Within-category	-2.58	0.24	35.39

Table 2 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 100-140. The across-category vowel change activated more strongly the left superior temporal gyrus (-59, -32, 8; *t*-value, -0.0025). The within-category vowel change activated more strongly the right superior temporal gyrus (60, -32, 15; *t*-value, -0.0021) (Table 2 and Fig. 2).

Table 2 Stereotaxic coordinates of the strongest activation foci during the across- and within-category-change discrimination

Vowel category		**Coordinates (mm)**			***t* values**
	BA	***x***	***y***	***z***	
Across-category					
L posterior superior temporal gyrus	22	-59	-32	8	-0.0025
Within-category					
R posterior superior temporal gyrus	22	60	-32	15	-0.0021

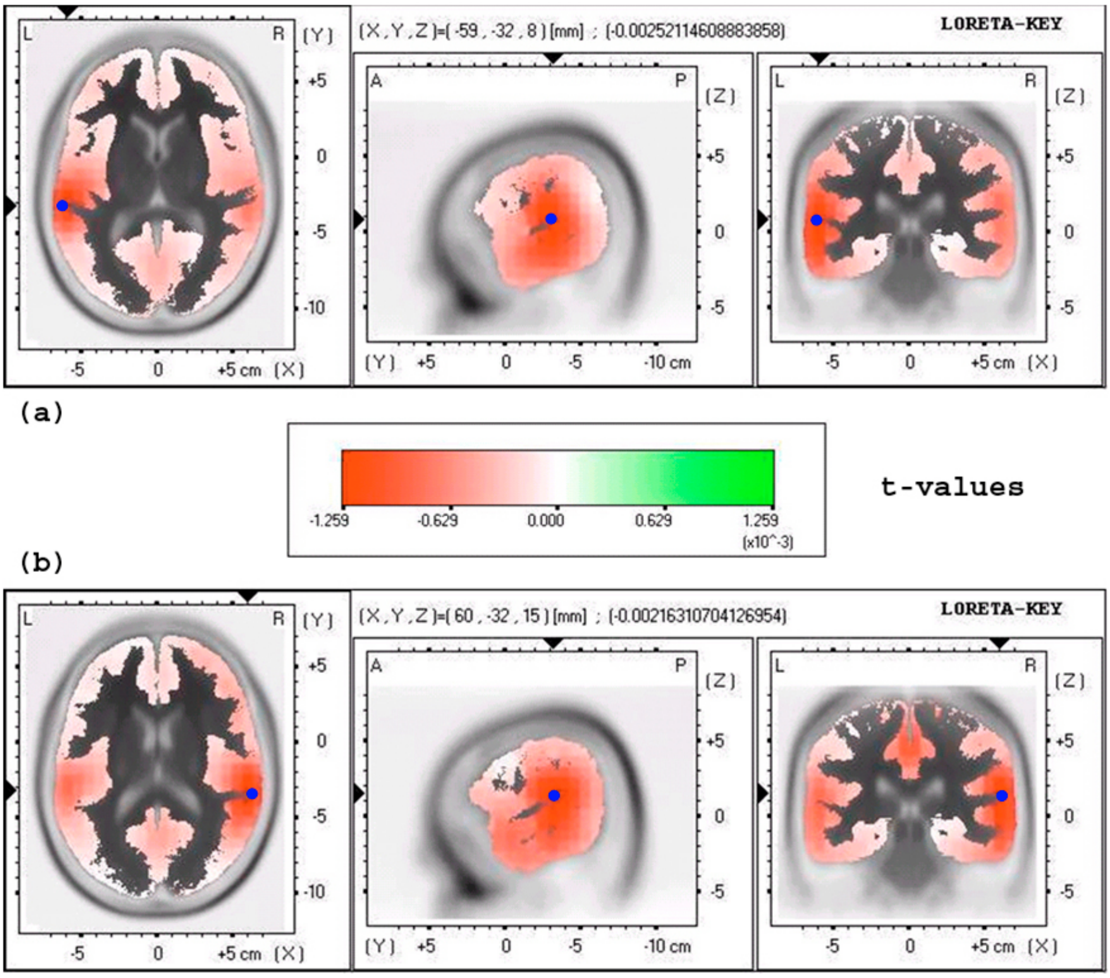

Figure 2: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the across-category (a) and the within-category (b) changes of vowels. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

4.1.2. Experiment 2: Preattentive detection of vowel duration changes with level tones

4.1.2.1. Subjects

Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 23-33 (mean 24.2; five females) and eleven NonS, aged 28-32 (mean 31.2; seven females). The approval of the institutional committee on human research and written consent from each subject were obtained (see appendix I).

4.1.2.2. Stimuli and procedure

Stimuli consisted of two pairs of monosyllabic, Thai words. Speech stimuli were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Pairs were designed to have similar mid tone. Two different stimuli were synthetically generated: (1) /k^{h}aam/ - long vowel, level tone and (2) /k^{h}am/ - short, level. The vowel-duration difference between (1) and (2) was 75 ms (628 vs. 553 ms) with the same level tone and intensity used in each stimulus. Five NS listened to the synthesized words and evaluated them all as natural sounding.

The standard (S)/deviant (D) pairs for each experiment, which was randomized across subjects, were shown [Experiment 1] S-1, D-2, [Experiment 2] S-2, D-1. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

4.1.2.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.1.2.4. Results

The grand-averaged ERPs show that both long-to-short and short-to-long duration changes with level tone perception elicited MMN between 184-208 ms with reference to the standard-stimulus ERPs. MMN amplitudes of the S and D was significantly generated a main effect of conditions, as indicated by the t-test comparison between the mean amplitude of MMN at Fz in Experiments 1 ($t_{(10)} = 4.95$, $P = 0.0006$). In Experiment 2, however, the S-D differences were not significant ($t_{(10)} = 0.37$, $P = 7169$, for the main effect of conditions). The result showed that long-to-short duration changes with level tone elicited a strong MMN bilaterally for NS and NonS, unlike short-to-long duration change with level tone (see Table 3). Furthermore, an across-experiment ANOVA demonstrated an interaction and main effects. The significant difference in MMN amplitudes was observed between groups across experiments ($F_{(3,30)} = 22.92$, $P < 0.0001$).

Table 3: Mean amplitude (μV) ±S.D. of MMN elicited by a vowel duration changes with level tone perception in NS and NonS.

Vowel duration change	Native (NS)	Non-native (NonS)
Long-to-short vowel changes	-0.57±0.16	-0.81±0.23
Short-to-long vowel changes	-0.52±0.16	-0.55±0.35

Source localization analyses were performed using LORETA-Key (Pascual-Marqui, 1994). Table 4 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 184-208 ms. In Experiment 1, a single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere for both groups. In Experiment 2, sources were obtained in the MTG and the superior temporal gyrus (STG) of the RH for both groups. The hemispheric difference was not discovered in this study (see Table 4 and Fig. 3- Fig. 6).

Table 4: Stereotaxic coordinates of activation foci during the vowel duration changes with level tone perception.

Vowel duration change		**Coordinates (mm)**			***t* values**
	BA	*x*	*y*	*z*	
Native speaker of Thai (NS)					
Long-to-short vowel change					
L middle temporal gyrus (MTG)	39	-52	-67	8	2.23
R middle temporal gyrus (MTG)	37	46	-67	8	1.44
Short-to-long vowel change					
L middle temporal gyrus (MTG)	39	-52	-67	8	1.26
R superior temporal gyrus (STG)	22	53	-60	15	1.14
Nonnative speaker of Thai (NonS)					
Long-to-short vowel change					
L middle temporal gyrus (MTG)	39	46	-67	8	2.52
R middle temporal gyrus (MTG)	37	46	-67	8	1.80
Short-to-long vowel change					
L superior temporal gyrus (STG)	42	-59	-32	15	1.23
R middle temporal gyrus (MTG)	21	60	-32	-6	1.16

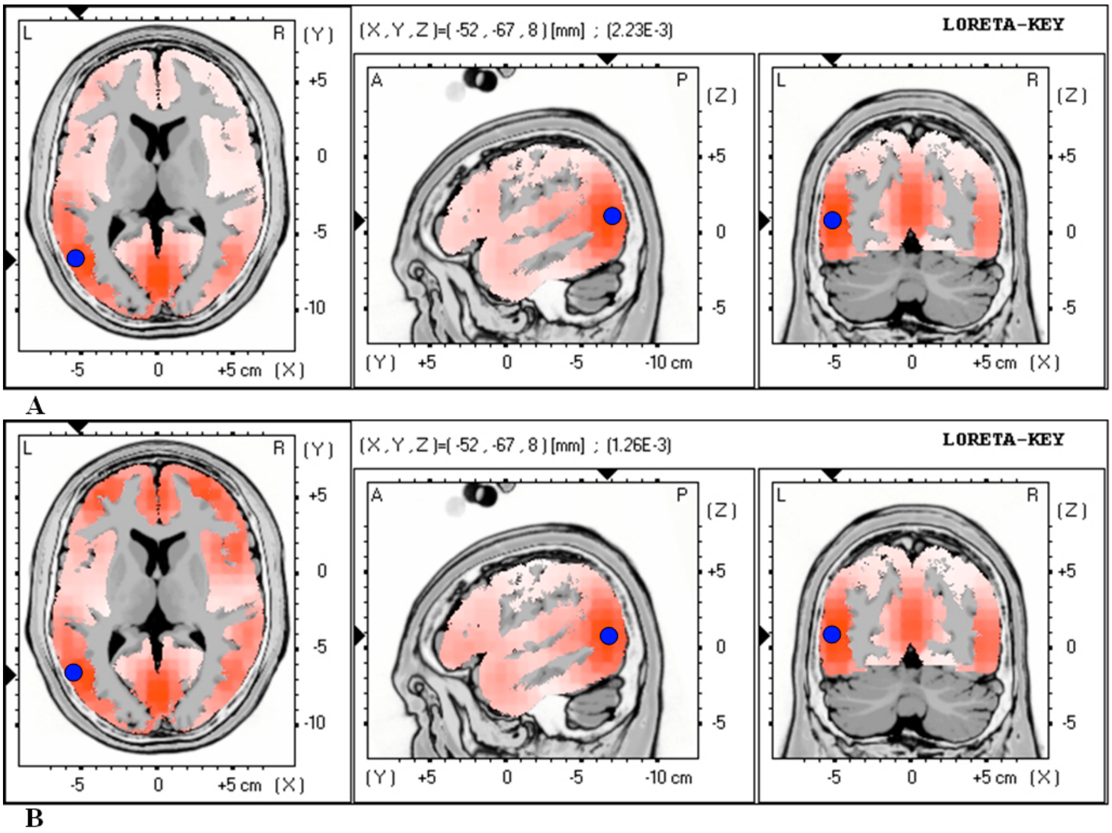

Figure 3: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with level tone of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

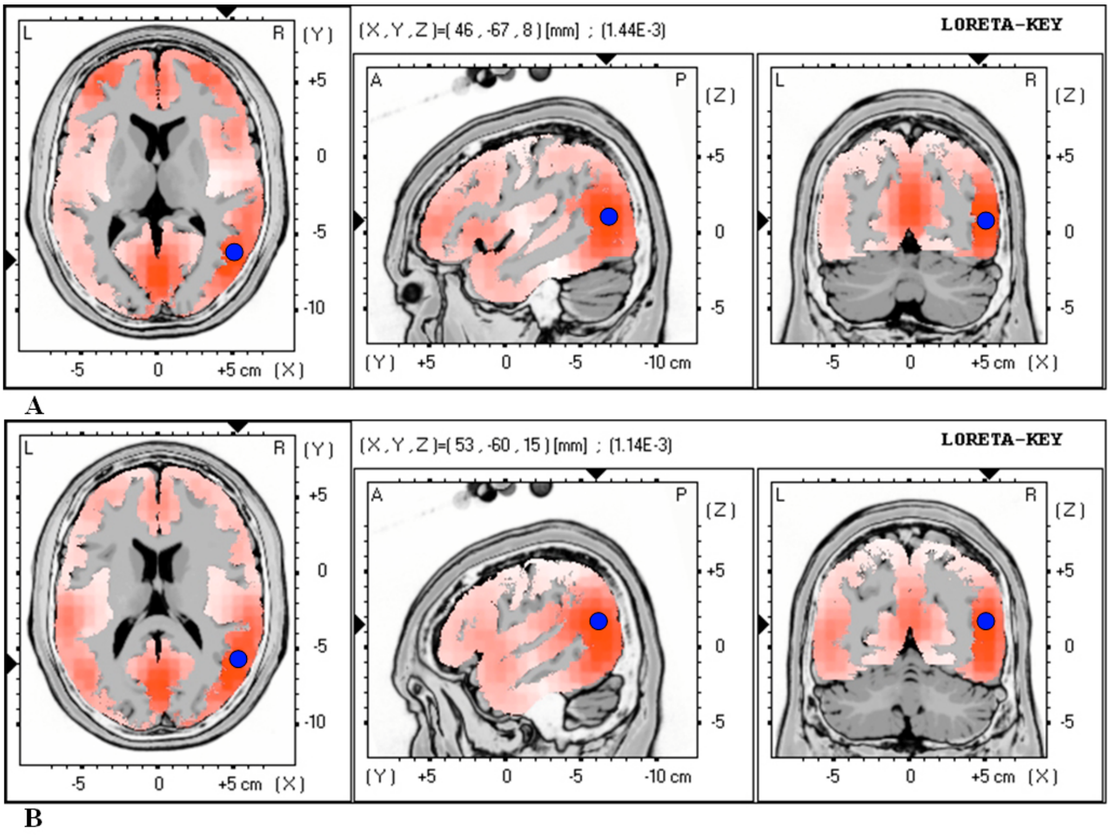

Figure 4: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with level ton of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

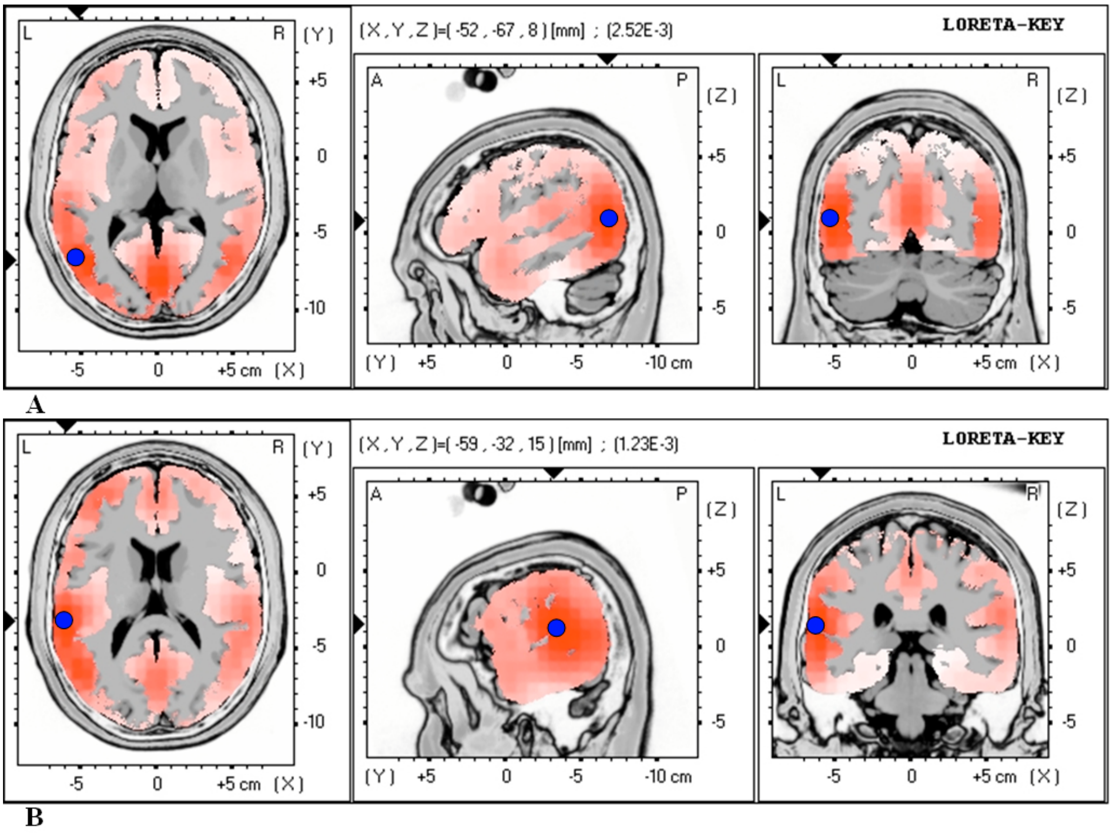

Figure 5: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with level tone of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

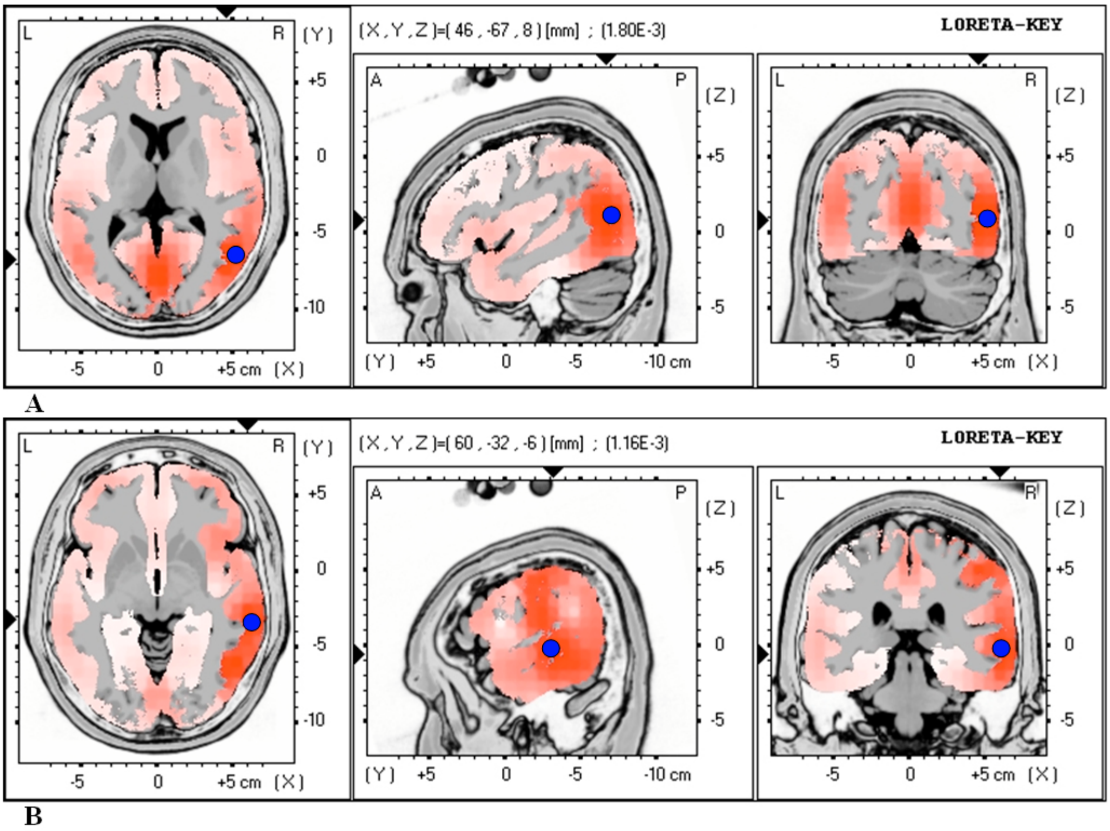

Figure 6: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with level tone of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

4.1.3. Experiment 3: Pre-attentive detection of vowel duration changes with contour tones

4.1.3.1. Subjects

Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 23-33 (mean 24.2; five females) and eleven NonS, aged 28-32 (mean 31.2; seven females). The approval of the institutional committee on human research and written consent from each subject were obtained (see appendix I).

4.1.3.2. Stimuli and procedure

Stimuli consisted of four pairs of monosyllabic, Thai words. Speech stimuli were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Four different stimuli were synthetically generated: (1) /k^{h}aam/ - long vowel, falling tone; (2) /k^{h}am/ - short vowel, falling; (3) /k^{h}aam/ - long, rising; (4) /k^{h}am/ - short, rising. The vowel-duration difference between (1) and (2) was 46 ms (546 vs. 500 ms) and between (3) and (4) was 56 ms (595 vs. 539 ms) with the same intensity used in each stimulus. Five NS listened to the synthesized words and evaluated them all as natural sounding.

The standard (S)/deviant (D) pairs for each experiment, which was randomized across subjects, were shown [Experiment 1] S-1, D-2, [Experiment 2] S-2, D-1, [Experiment 3] S-3, D4, [Experiment 4] S-4, D-3. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

4.1.3.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.1.3.4. Results

The grand-averaged ERPs show that both rising-to-falling and falling-to-rising tone changes perception elicited MMN between 172-264 ms with reference to the standard-stimulus ERPs. An ANOVA comparing MMN amplitudes of the S and D yield a main effect of conditions in Experiments 1 ($F_{(3,40)} = 8.61$, $P < 0.0001$) and 3 ($F_{(3,40)} = 23.62$, $P < 0.0001$). In Experiment 2 and 4, however, the S-D differences were not significant (e.g., $F_{(3,40)} = 1.22$, $P = 0.2511$ in Experiment 2, n.s.; $F_{(3,40)} = 0.52$, $P = 0.615$ in Experiment 4, n.s., for the main effect of conditions). The result showed that long-to-short duration with falling and rising tone changes elicited a strong MMN bilaterally for NS and NonS, unlike short-to-long duration with falling and rising tone changes (see Table 5). Furthermore, an across-experiment ANOVA demonstrated an interaction and main effects. The significant difference in MMN amplitudes was observed between groups across experiments ($F_{(7,80)} = 45.61$, $P < 0.0001$).

Table 5: Mean amplitude (μV)±S.D. of MMN elicited by a vowel duration changes with contour tones perception in NS and NonS.

	Native (NS)	**Nonnative (NonS)**
Long (falling)-to-Short (falling) changes	-1.61±0.18	-1.12±0.59
Short (falling)-to-Long (falling) changes	-1.16±0.24	-1.09±0.24
Long (rising)-to-Short (rising) changes	-1.22±0.31	-0.30±0.21
Short (rising)-to-Long (rising) changes	-0.93±0.25	-0.11±0.49

Source localization analyses were performed using LORETA-Key (Pascual-Marqui, 1994). Table 6 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 172-264 ms. In Experiment 2 and 3, a single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere for both groups. In Experiment 1 and 4, sources were obtained in the MTG of the LH and in the superior temporal gyrus (STG) of the RH for both groups. No hemispheric difference was discovered in this study (see Table 6 and Fig. 7- Fig. 14).

4.2. Auditory cortex responses to vowel duration and pitch changes

4.2.1. Experiment 4: Preattentive detection of level-to-contour tone changes

4.2.1.1. Subjects

Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 23-33 (mean 24.2; five females) and eleven NonS, aged 28-32 (mean 31.2; seven females). The approval of the institutional committee on human research and written consent from each subject were obtained (see appendix I).

4.2.1.2. Stimuli and procedure

Stimuli consisted of four pairs of monosyllabic, Thai words. Speech stimuli were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Pairs were designed to have similar long vowel duration. Three different stimuli were synthetically generated: (1) /k^{h}aam/ - level tone; (2) /k^{h}aam/ - falling; (3) /k^{h}aam/ - rising. Five NS listened to the synthesized words and evaluated them all as natural sounding.

Table 6: Stereotaxic coordinates of activation foci during the vowel duration changes with contour tones perception.

Vowel duration change		**Coordinates (mm)**			***t* values**
	BA	***x***	***y***	***z***	
Native speaker of Thai (NS)					
Long (falling)-to-short (falling) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	4.15
R superior temporal gyrus (STG)	22	53	-60	15	2.69
Short (falling)-to-long (falling) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	3.58
R superior temporal gyrus (STG)	22	53	-60	15	1.89
Long (rising)-to-short (rising) change					
L middle temporal gyrus (MTG)	21	-52	3	-20	3.64
R middle temporal gyrus (MTG)	37	46	-67	8	2.79
Short (rising)-to-long (rising) change					
L middle temporal gyrus (MTG)	21	-52	3	-20	2.39
R superior temporal gyrus (STG)	22	53	-60	15	2.06
Nonnative speaker of Thai (NonS)					
Long (falling)-to-short (falling) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	3.60
R superior temporal gyrus (STG)	22	53	-60	15	2.59
Short (falling)-to-long (falling) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	2.73
R middle temporal gyrus (MTG)	37	46	-67	8	1.59
Long (rising)-to-short (rising) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	3.38
R superior temporal gyrus (STG)	22	53	-60	15	1.65
Short (rising)-to-long (rising) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	2.81
R superior temporal gyrus (STG)	21	60	-32	-6	2.12

The standard (S)/deviant (D) pairs for each experiment, which was randomized across subjects, were shown [Experiment 1] S-1, D-2, [Experiment 2] S-2, D-1, [Experiment 3] S-1, D-3, [Experiment 4] S-3, D-1. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

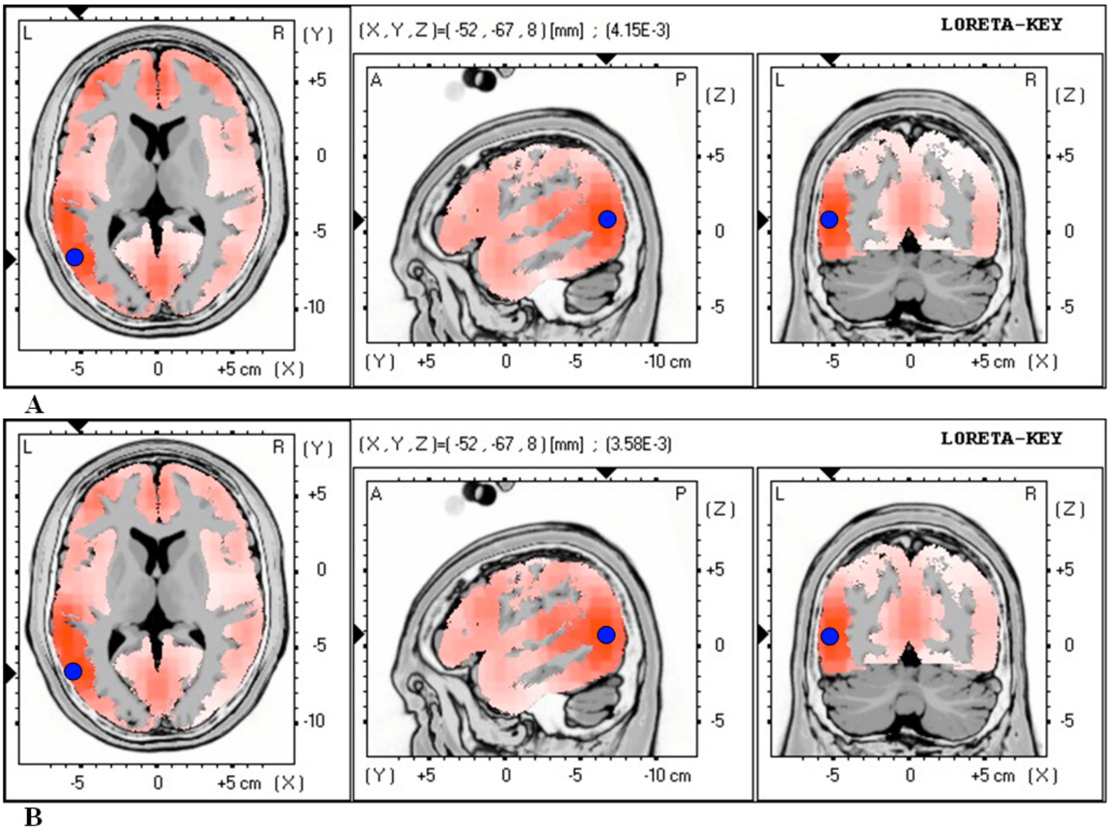

Figure 7: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with falling tone of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

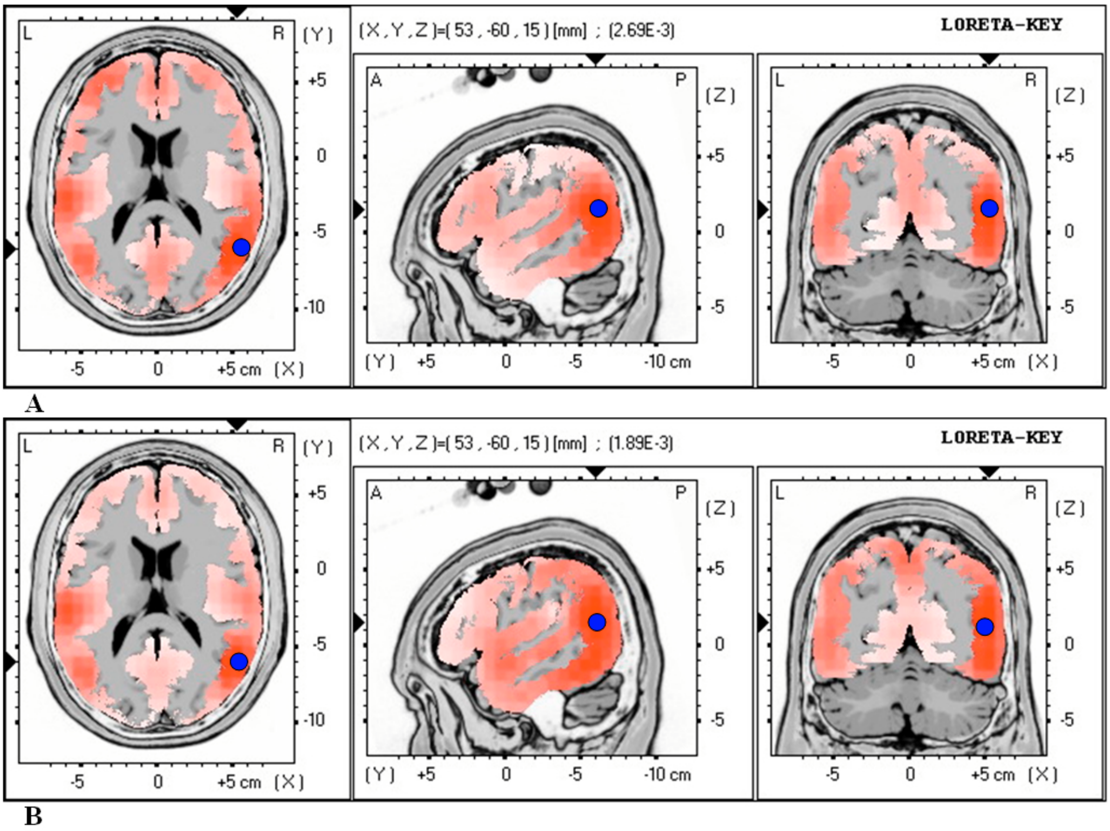

Figure 8: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with falling tone of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

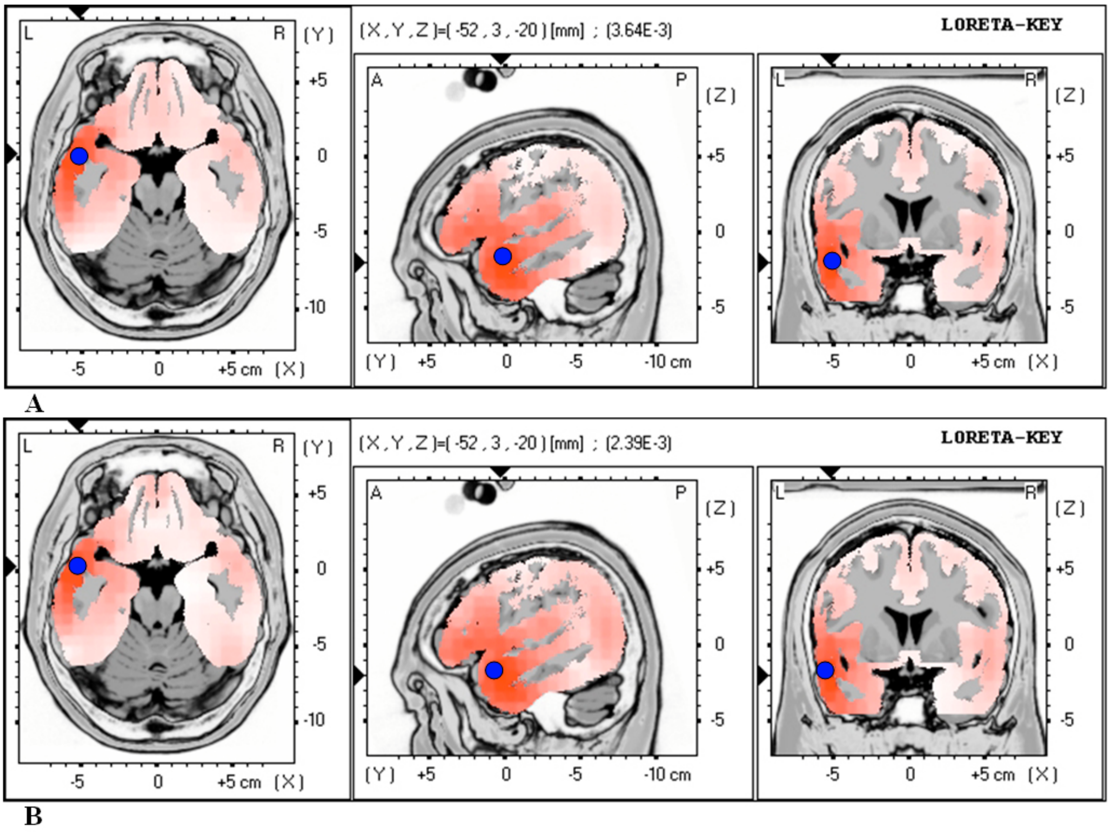

Figure 9: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with rising tone of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

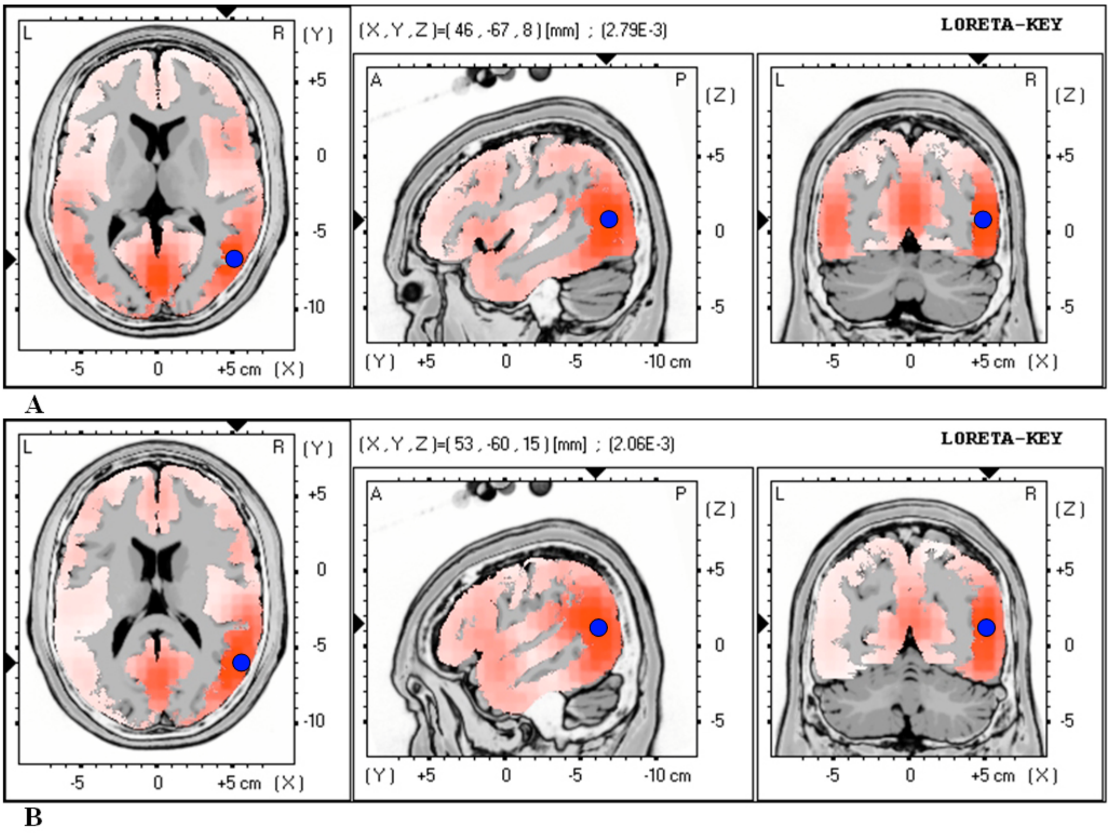

Figure 10: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with rising tone of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

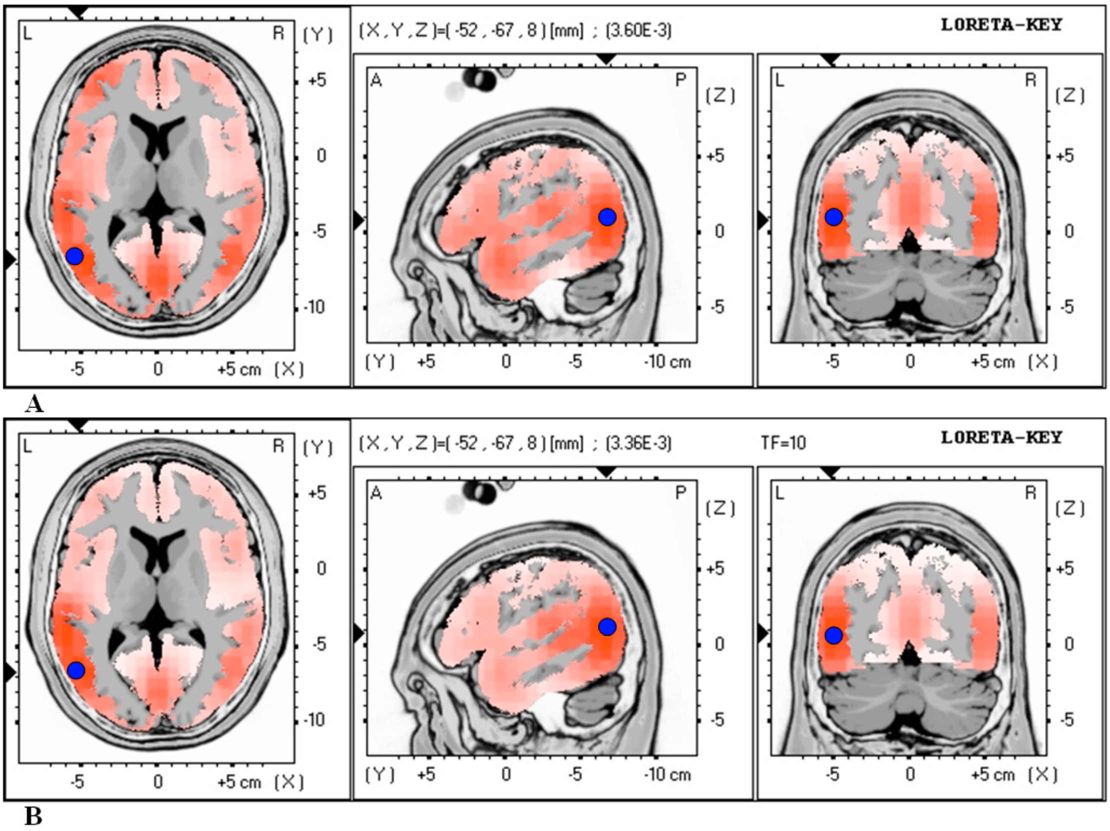

Figure 11: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with falling tone of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

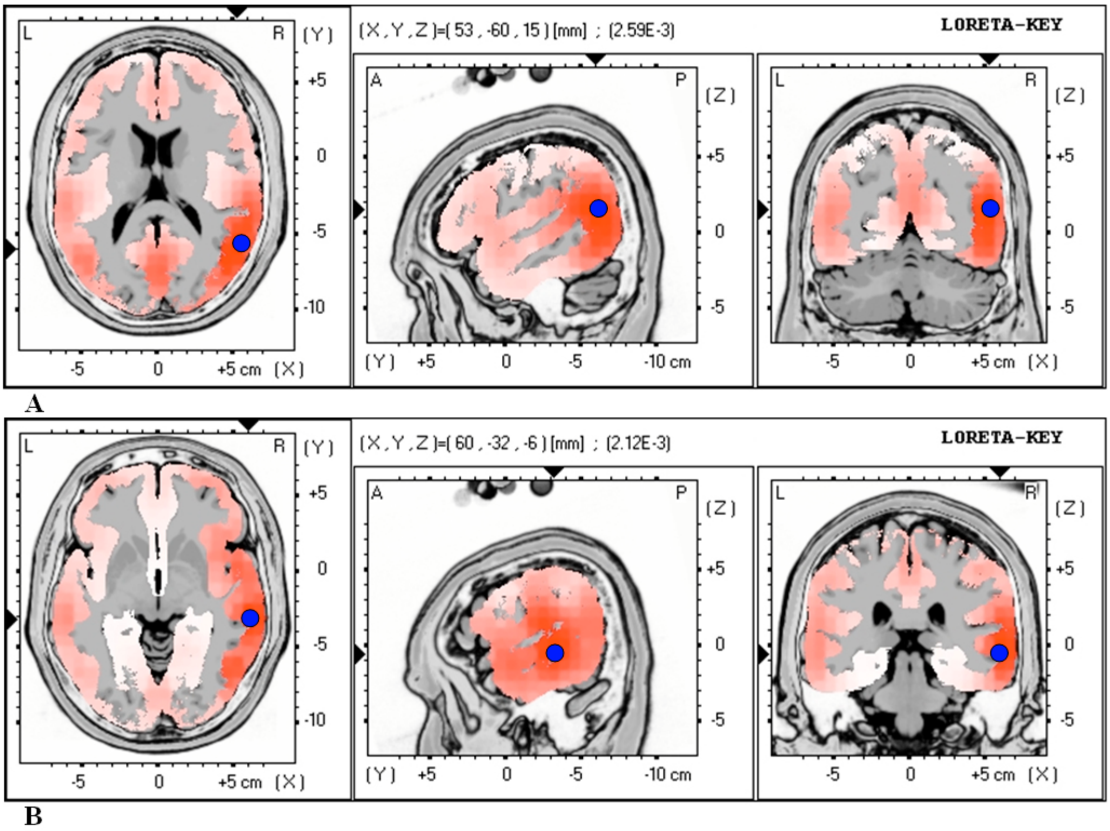

Figure 12: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with falling tone of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

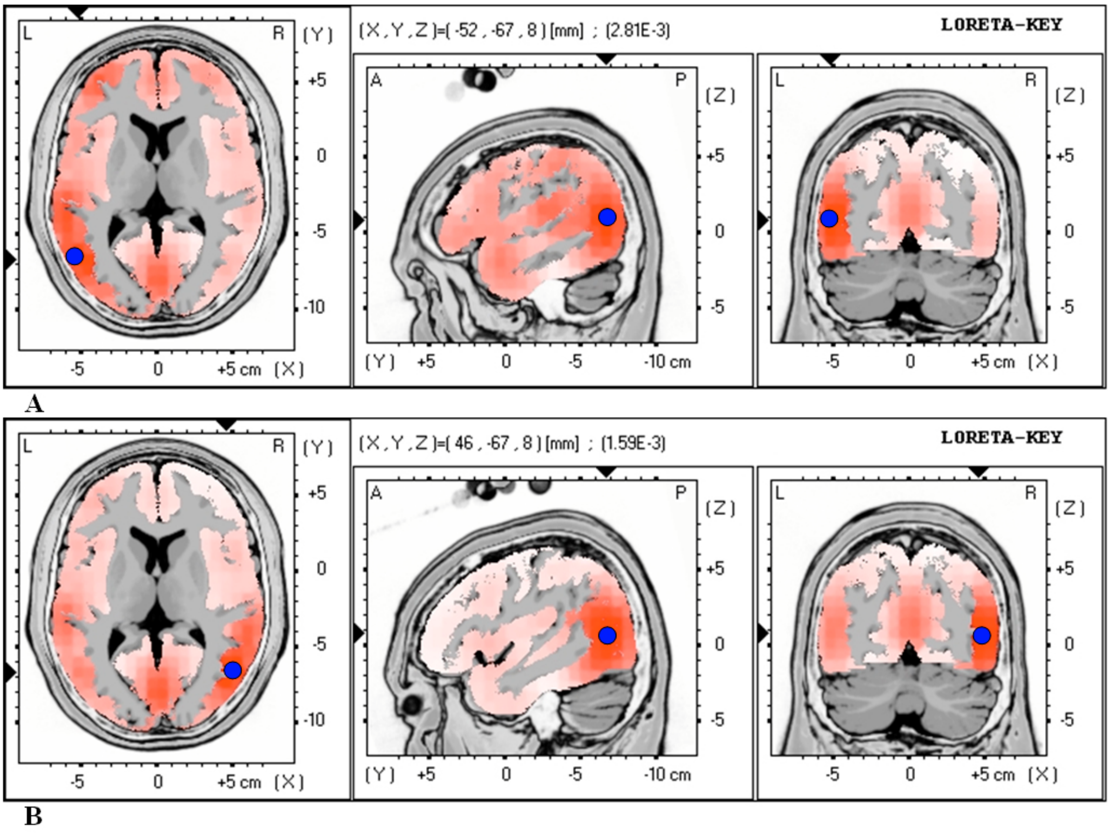

Figure 13: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with rising tone of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

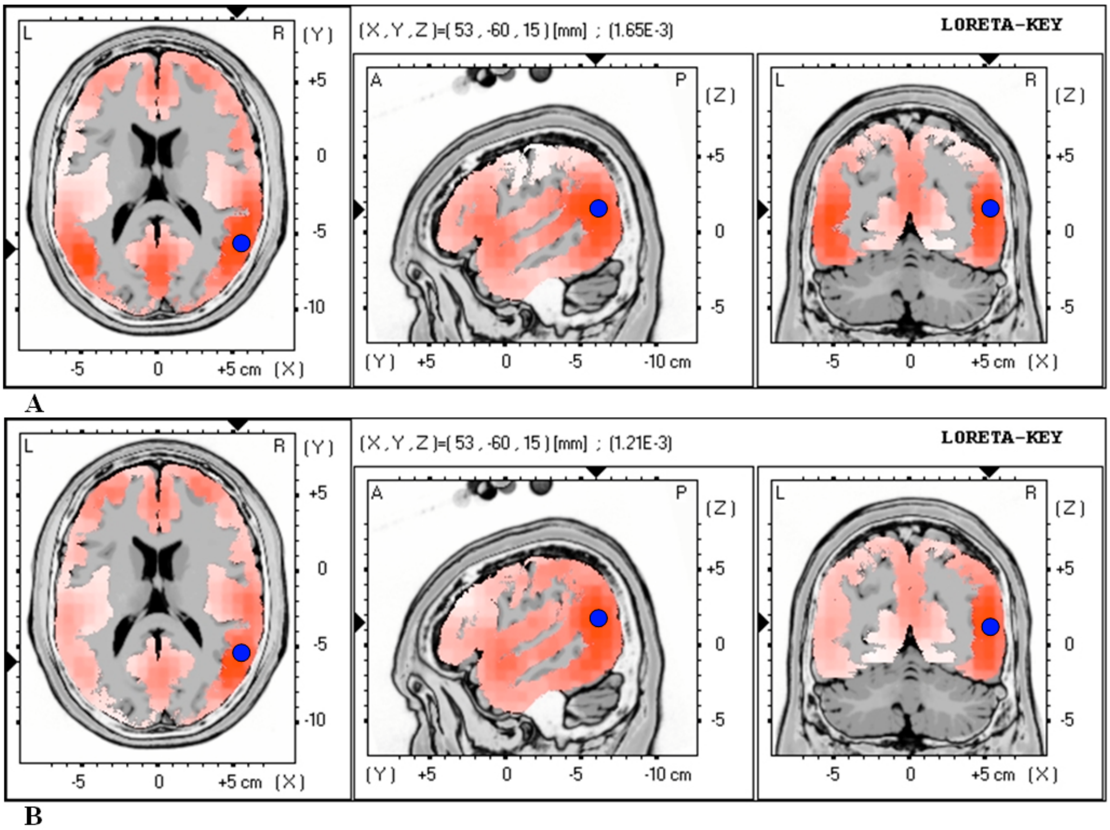

Figure 14: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short duration (A) and the short-to-long duration (B) changes of vowels with rising tone of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

4.2.1.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.2.1.4. Results

The grand-averaged ERPs show that both level-to-contour and contour-to-level tone changes perception elicited MMN between 196-220 ms with reference to the standard-stimulus ERPs. An ANOVA comparing MMN amplitudes of the S and D yield a main effect of conditions in Experiments 1 ($F_{(3,40)} = 10.65$, $P < 0.0001$) and 3 ($F_{(3,40)} = 6.64$, $P < 0.0001$). In Experiment 2 and 4, however, the S-D differences were not significant (e.g., $F_{(3,40)} = 1.44$, $P = 0.1808$ in Experiment 2, n.s.; $F_{(3,40)} = 1.11$, $P = 0.2948$ in Experiment 4, n.s., for the main effect of conditions). The result showed that level-to-falling and level-to-rising tone changes elicited a strong MMN bilaterally for NS and NonS, unlike falling-to-level and rising-to-level tone changes (see Table 7). Furthermore, an across-experiment ANOVA demonstrated an interaction and main effects. The significant difference in MMN amplitudes was observed between groups across experiments ($F_{(7,80)} = 46.66$, $P < 0.0001$).

Table 7: Mean amplitude (µV)±S.D. of MMN elicited by a level-to-contour tones perception in NS and NonS.

Tone changes	Native (NS)	Non-native (NonS)
Mid-to-falling tone changes	-3.03±0.74	-2.41±0.94
Falling-to-mid tone changes	-0.40±0.55	-2.77±0.62
Mid-to-rising tone changes	-1.62±0.42	-2.22±0.59
Rising-to-mid tone changes	-0.20±0.41	-1.58±0.56

Source localization analyses were performed using LORETA-Key (Pascual-Marqui, 1994). Table 8 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 196-220 ms. In all Experiments, a single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere for NonNP group. In Experiment 1, 3 and 4, a single source was estimated to be located in the middle temporal gyrus (MTG), for Experiment 4 and the superior temporal gyrus (STG), for Experiment 1 and 3, of each hemisphere for NS group. In Experiment 2, sources were obtained in the MTG of the LH and in the superior temporal gyrus (STG) of the RH for NonS group. (see Table 8 and Fig. 15- Fig. 22).

Table 8: Stereotaxic coordinates of activation foci during the level-to-contour tone changes perception.

Tone change	**BA**	**Coordinates (mm)** *x*	*y*	*z*	***t* values**
Native speaker of Thai (NS)					
Mid-to-Falling tone change					
L superior temporal gyrus (STG)	22	-52	-60	15	3.73
R superior temporal gyrus (STG)	22	53	-60	15	3.42
Falling-to-Mid tone change					
L middle temporal gyrus (MTG)	39	-52	-67	8	3.48
R superior temporal gyrus (STG)	22	60	-39	15	2.12
Mid-to-Rising tone change					
L superior temporal gyrus (STG)	42	-59	-32	8	3.67
R superior temporal gyrus (STG)	39	53	-53	15	3.20
Rising-to-Mid tone change					
L middle temporal gyrus (MTG)	39	-52	-67	8	2.59
R middle temporal gyrus (MTG)	37	46	-67	8	1.11
Non-native speaker of Thai (NonS)					
Mid-to-Falling tone change					
L middle temporal gyrus (MTG)	39	-52	-67	8	4.58
R middle temporal gyrus (MTG)	39	53	-60	8	3.90
Falling-to-Mid tone change					
L middle temporal gyrus (MTG)	39	-52	-67	8	4.52
R middle temporal gyrus (MTG)	21	60	-32	-6	4.44
Mid-to-Rising tone change					
L middle temporal gyrus (MTG)	21	-59	-32	1	4.40
R middle temporal gyrus (MTG)	37	46	-67	8	1.06
Rising-to-Mid tone change					
L middle temporal gyrus (MTG)	39	-52	-67	8	4.28
R middle temporal gyrus (MTG)	39	53	-60	8	3.83

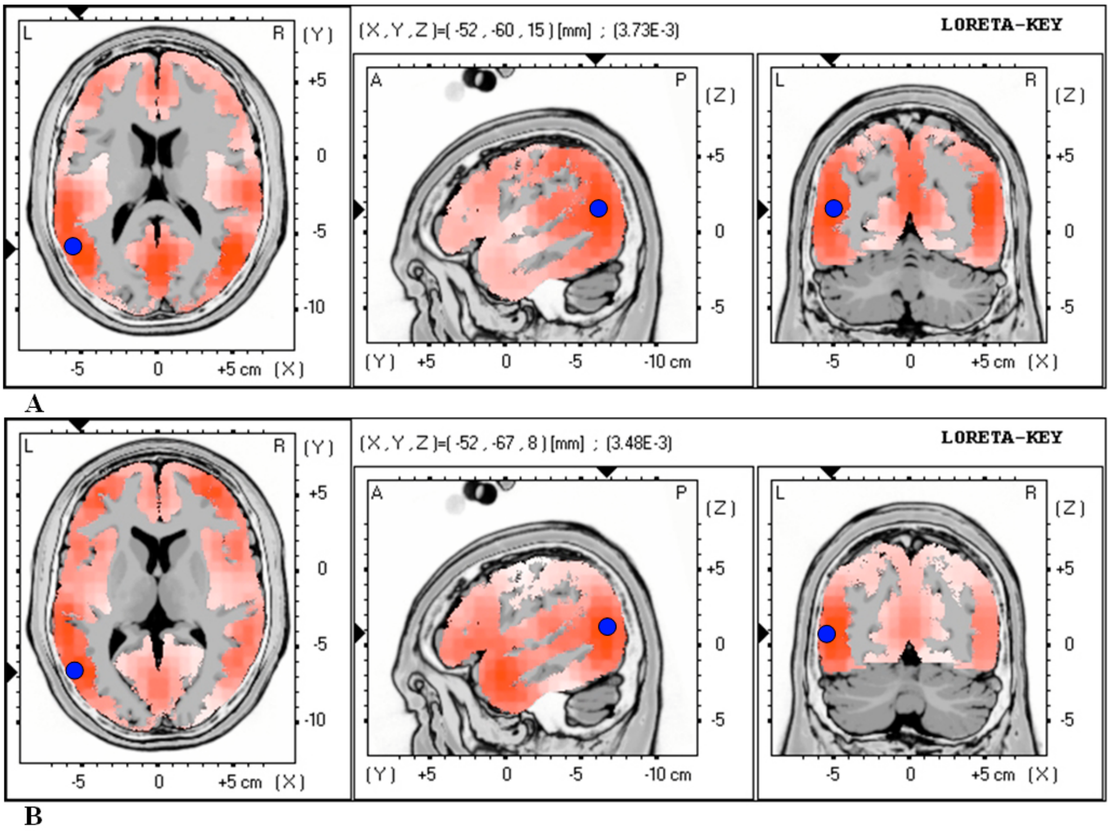

Figure 15: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-falling (A) and the falling-to-level (B) changes of tones with long vowel duration of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

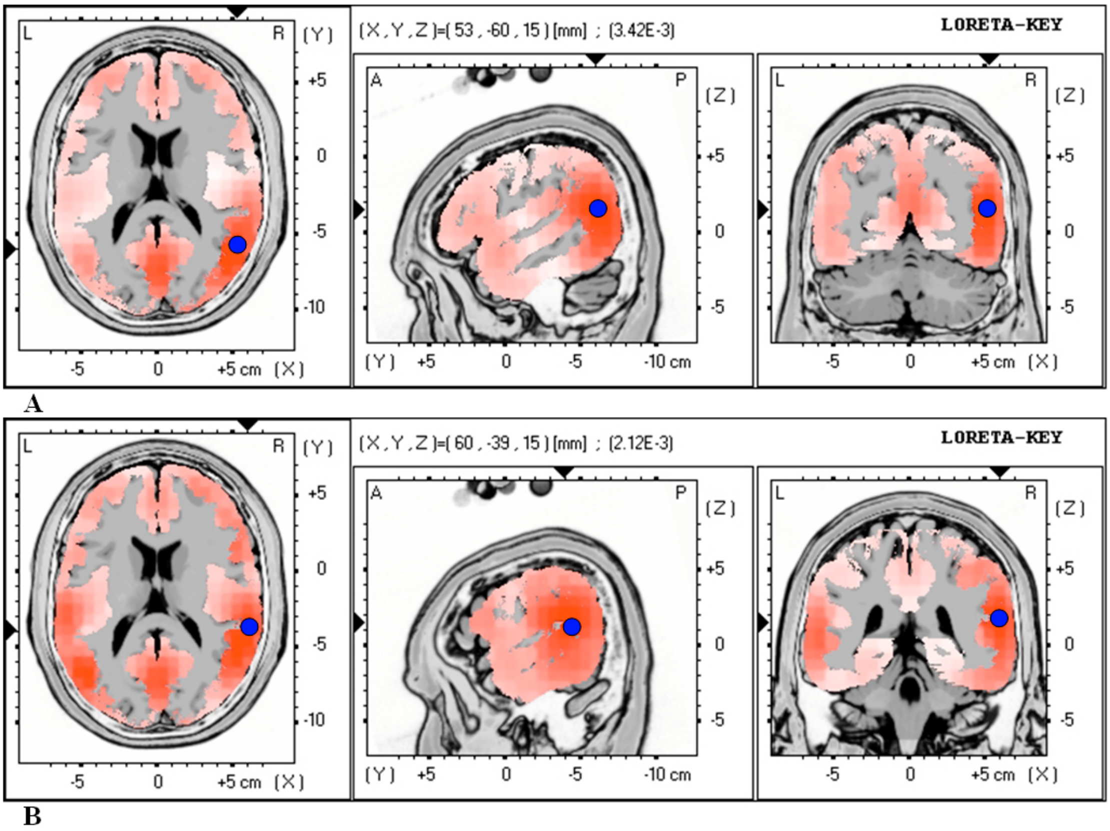

Figure 16: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-falling (A) and the falling-to-level (B) changes of tones with long vowel duration of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

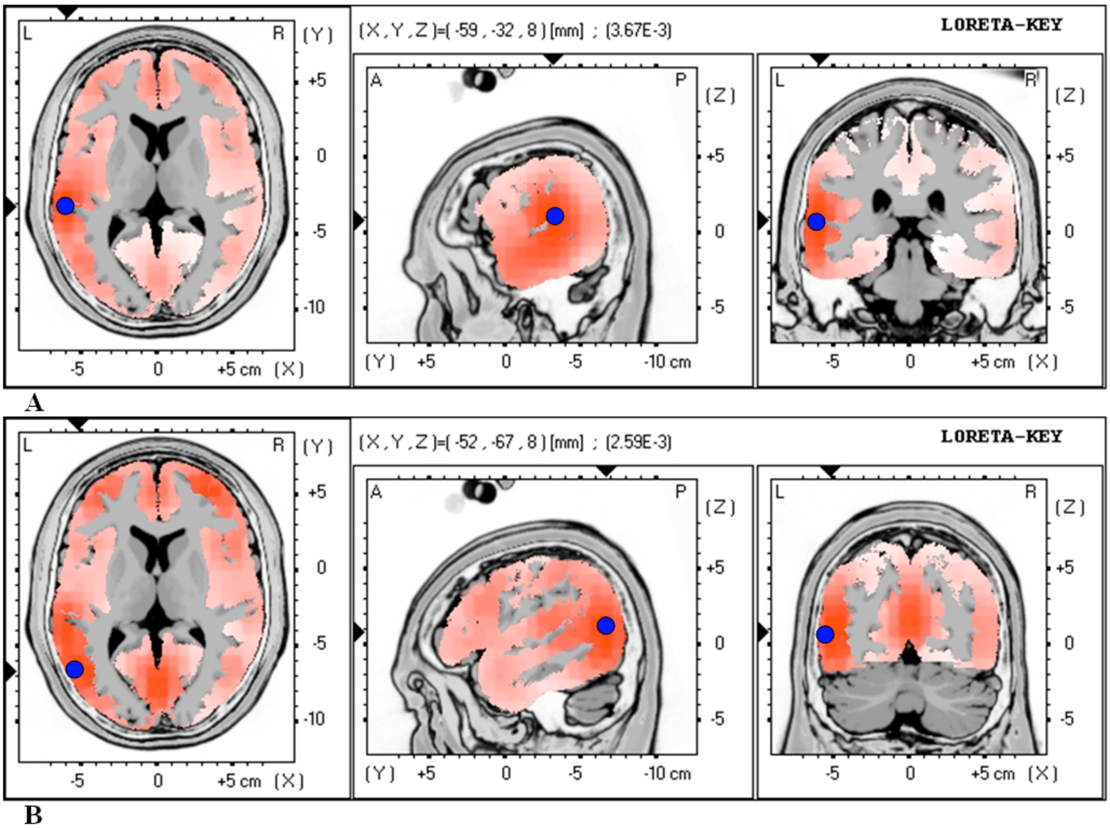

Figure 17: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-rising (A) and the rising-to-level (B) changes of tones with long vowel duration of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

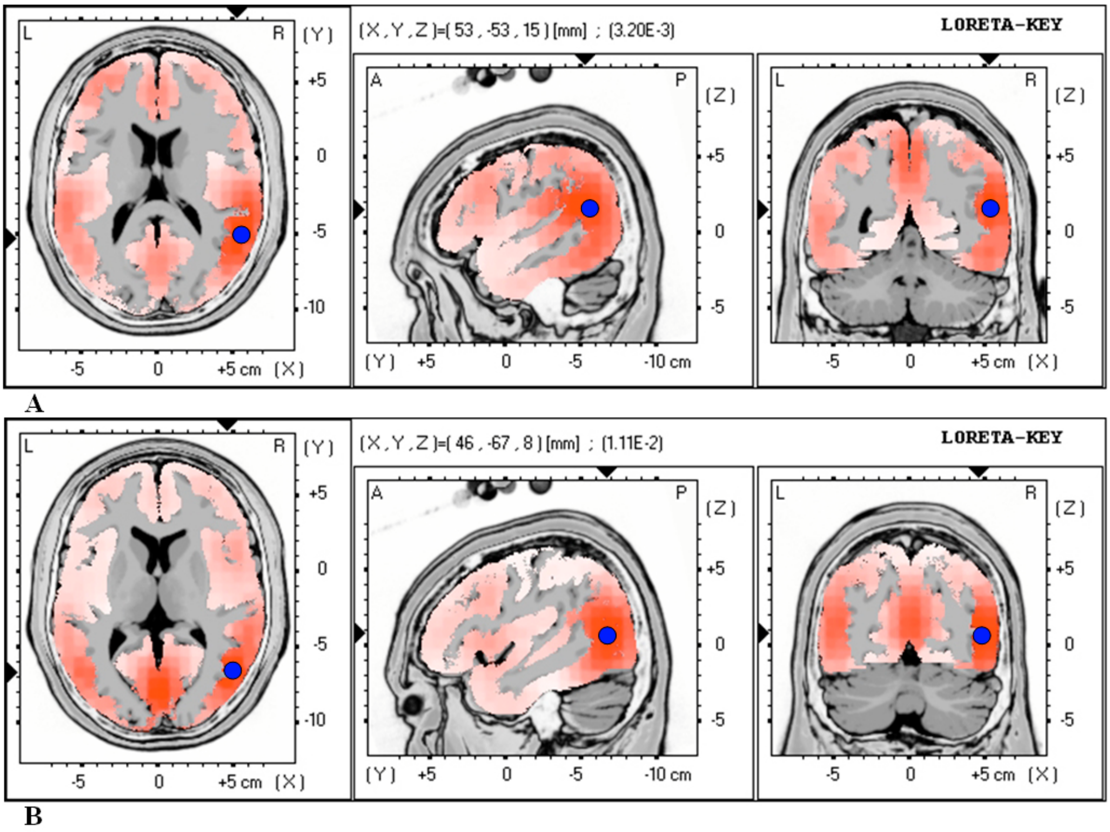

Figure 18: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-rising (A) and the rising-to-level (B) changes of tones with long vowel duration of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

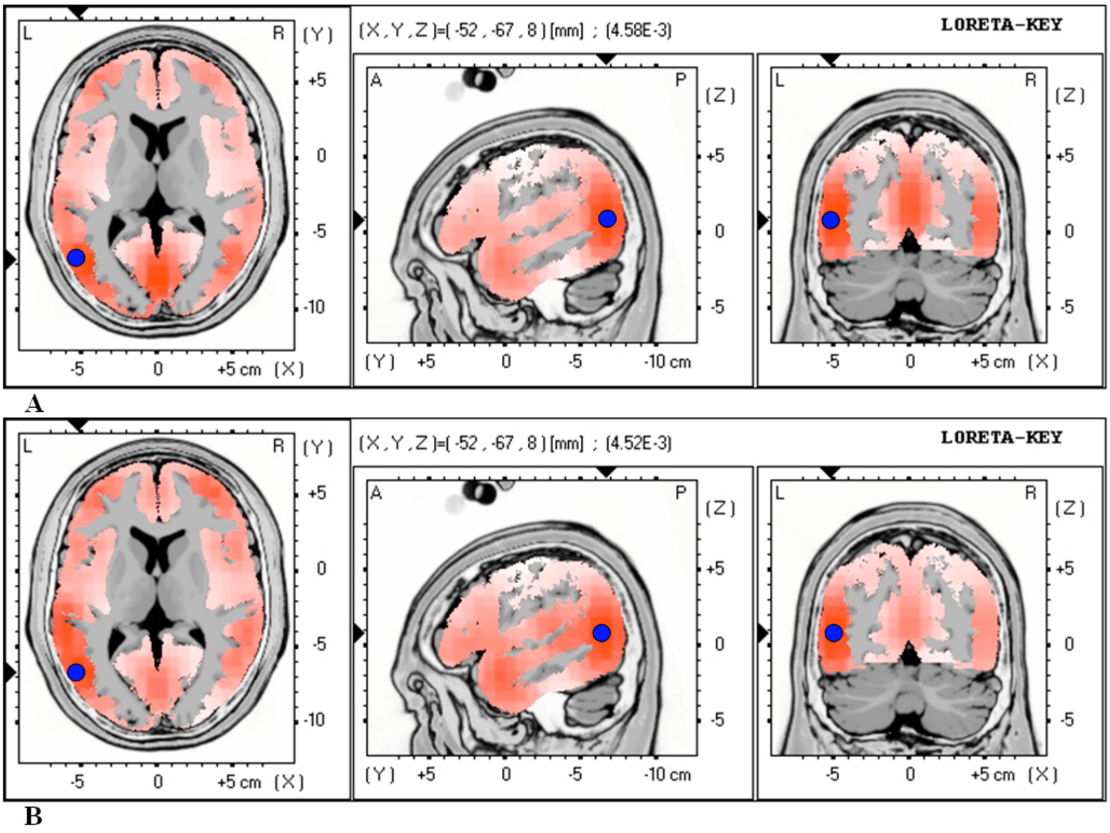

Figure 19: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-falling (A) and the falling-to-level (B) changes of tones with long vowel duration of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

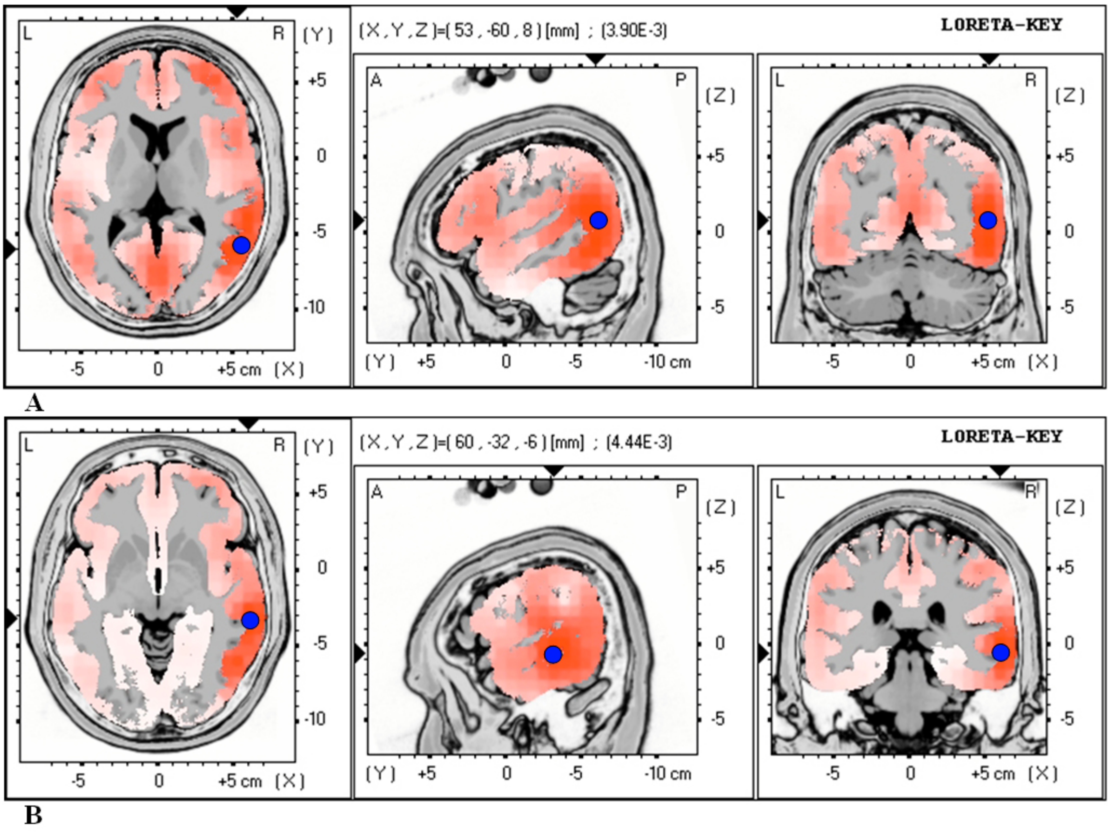

Figure 20: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-falling (A) and the falling-to-level (B) changes of tones with long vowel duration of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

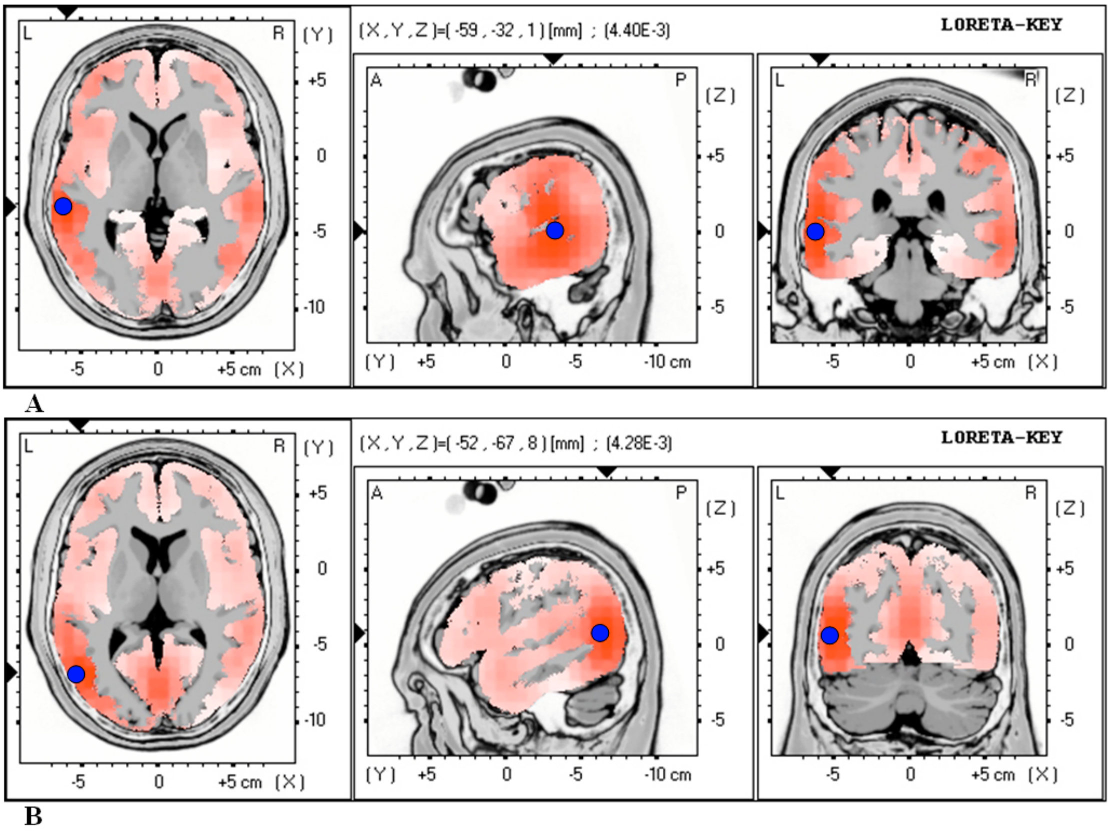

Figure 21: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-rising (A) and the rising-to-level (B) changes of tones with long vowel duration of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

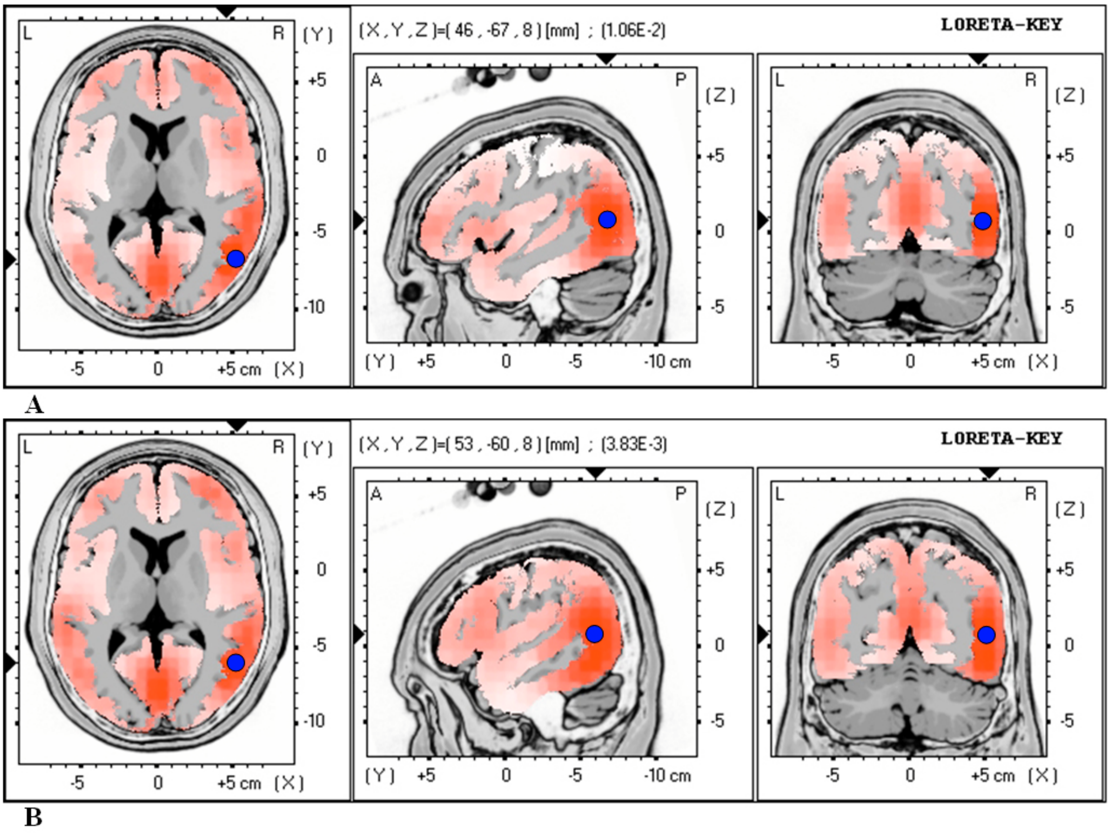

Figure 22: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the level-to-rising (A) and the rising-to-level (B) changes of tones with long vowel duration of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

4.2.2. Experiment 5: Preattentive detection of contour tone changes

4.2.2.1. Subjects

Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 23-39 (mean 25.3; six females) and eleven NonS, aged 23-29 (mean 27.8; nine females). The approval of the institutional committee on human research and written consent from each subject were obtained (see appendix I).

4.2.2.2. Stimuli and procedure

Stimuli consisted of two pairs of monosyllabic, Thai words. Speech stimuli were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Pairs were designed to have similar long vowel duration. Two different stimuli were synthetically generated: (1) /k^{h}aam/ - falling tone; (2) /k^{h}aam/ - rising. Five NS listened to the synthesized words and evaluated them all as natural sounding.

The standard (S)/deviant (D) pairs for each experiment, which was randomized across subjects, were shown [Experiment 1] S-2, D-1, [Experiment 2] S-1, D-2. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

4.2.2.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.2.2.4. Results

The grand-averaged ERPs show that both rising-to-falling and falling-to-rising tone changes perception elicited MMN between 212-244 ms with reference to the standard-stimulus ERPs. An ANOVA comparing MMN amplitudes of the S and D were not significant (e.g., $F_{(3,40)} = .62$, $P = 0.55$ in Experiment 1, n.s. and $F_{(3,40)} = 0.23$, $P = 0.82$ in Experiment 2, n.s., for the main effect of conditions). The result showed that rising-to-falling and falling-to-rising tone changes elicited a strong MMN bilaterally for NS and NonS (see Table 9). Furthermore, an across-experiment ANOVA demonstrated no interaction and main effects. The significant difference in MMN amplitudes was not observed between groups across experiments ($F_{(7,80)} = 0.12$, $P = 0.9555$, n.s.).

Table 9: Mean amplitude (µV)±S.D. of MMN elicited by a contour tone change perception in NS and NonS.

Tone changes	Native (NS)	Nonnative (NonS)
Rising-to-falling tone changes	-3.03±0.87	-3.13±1.08
Falling-to-rising tone changes	-2.92±0.95	-2.99±0.70

Source localization analyses were performed using LORETA-Key (Pascual-Marqui, 1994). Table 10 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 212-224 ms. In Experiment 2, a single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere for both groups. In Experiment 1, sources were obtained in the MTG of the LH and in the superior temporal gyrus (STG) of the RH for both groups. No hemispheric difference was discovered in this study (see Table 10 and Fig. 23- Fig. 26).

Table 10: Stereotaxic coordinates of activation foci during the contour tones change perception.

Tone change	**BA**	**Coordinates (mm)** *x*	 *y*	 *z*	***t* values**
Native speakers of Thai (NS)					
Rising-to-falling of tone change					
L middle temporal gyrus (MTG)	39	-45	-67	15	5.91
R superior temporal gyrus (STG)	22	53	-60	15	4.69
Falling-to-rising of tone change					
L middle temporal gyrus (MTG)	21	-59	-32	1	5.73
R middle temporal gyrus (MTG)	37	46	-67	8	4.64
Nonnative speaker of Thai (NonS)					
Rising-to-falling of tone change					
L middle temporal gyrus (MTG)	39	-52	-67	15	1.82
Right superior temporal gyrus (STG)	22	53	-60	15	1.12
Falling-to-rising of tone change					
L middle temporal gyrus (MTG)	39	-45	-67	15	1.10
R middle temporal gyrus (MTG)	37	46	-67	8	1.06

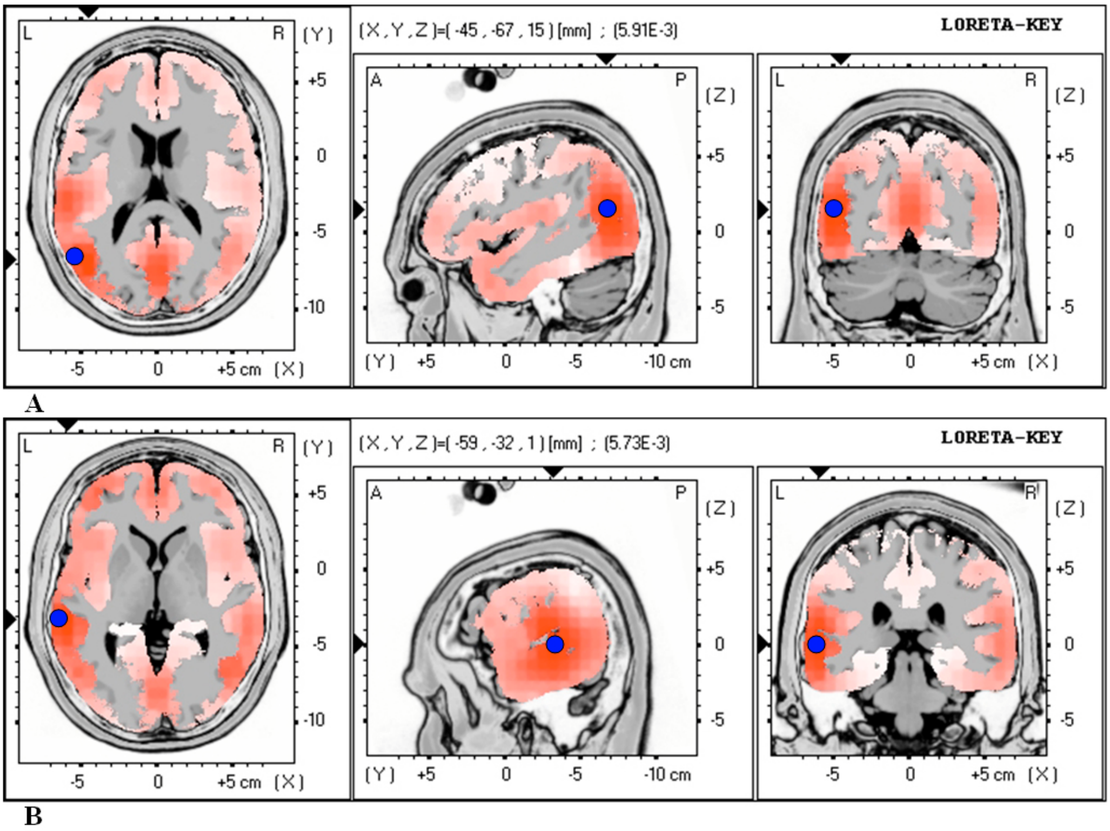

Figure 23: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the rising-to-falling (A) and the falling-to-rising (B) changes of tones with long vowel duration of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

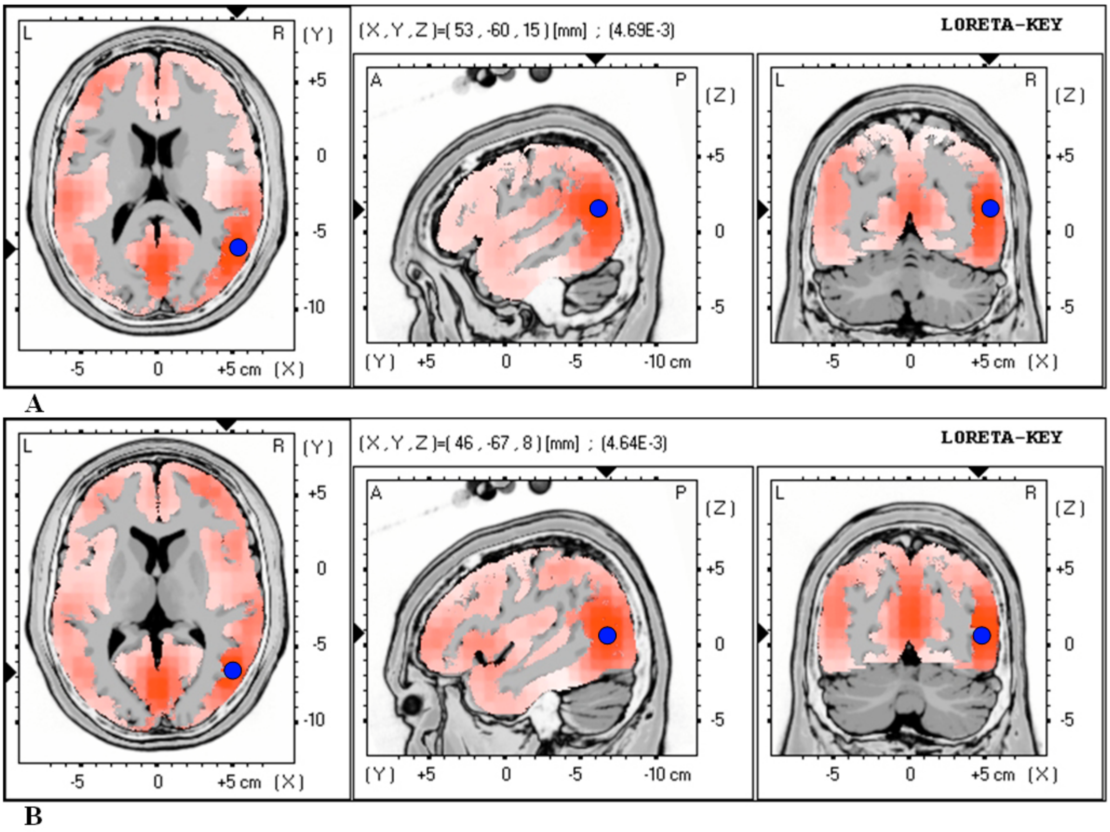

Figure 24: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the rising-to-falling (A) and the falling-to-rising (B) changes of tones with long vowel duration of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

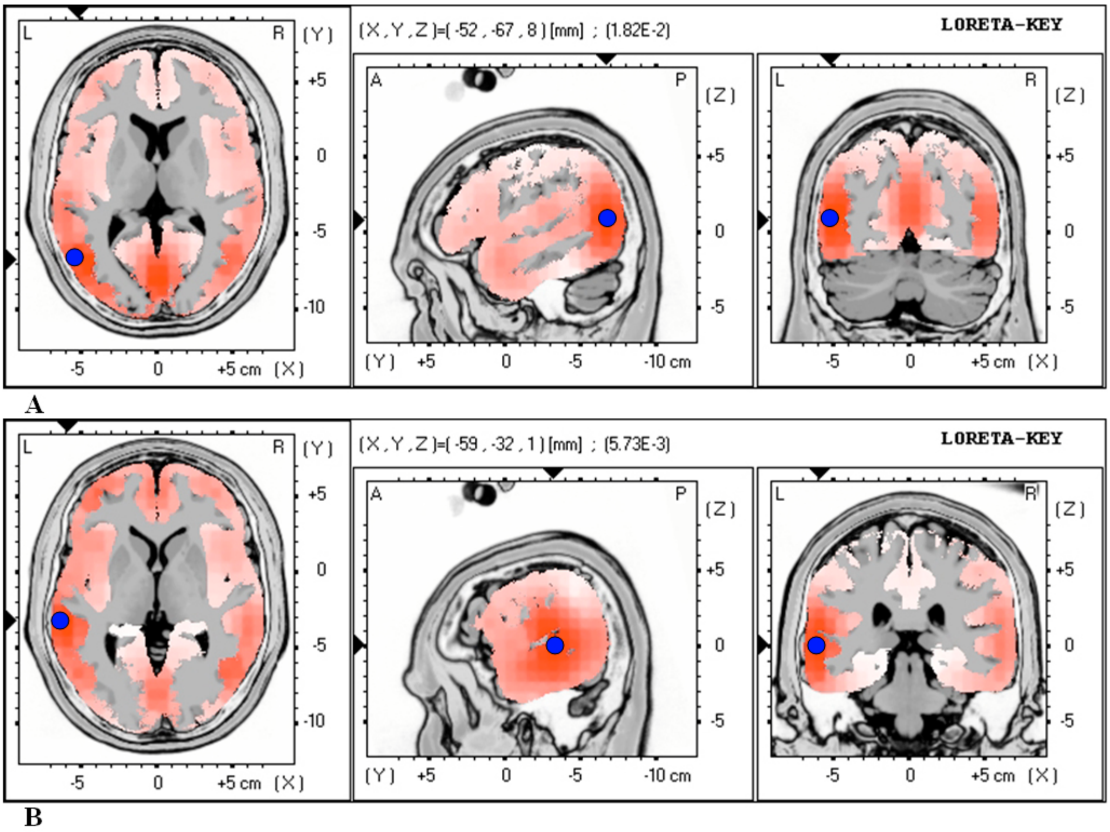

Figure 25: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the rising-to-falling (A) and the falling-to-rising (B) changes of tones with long vowel duration of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

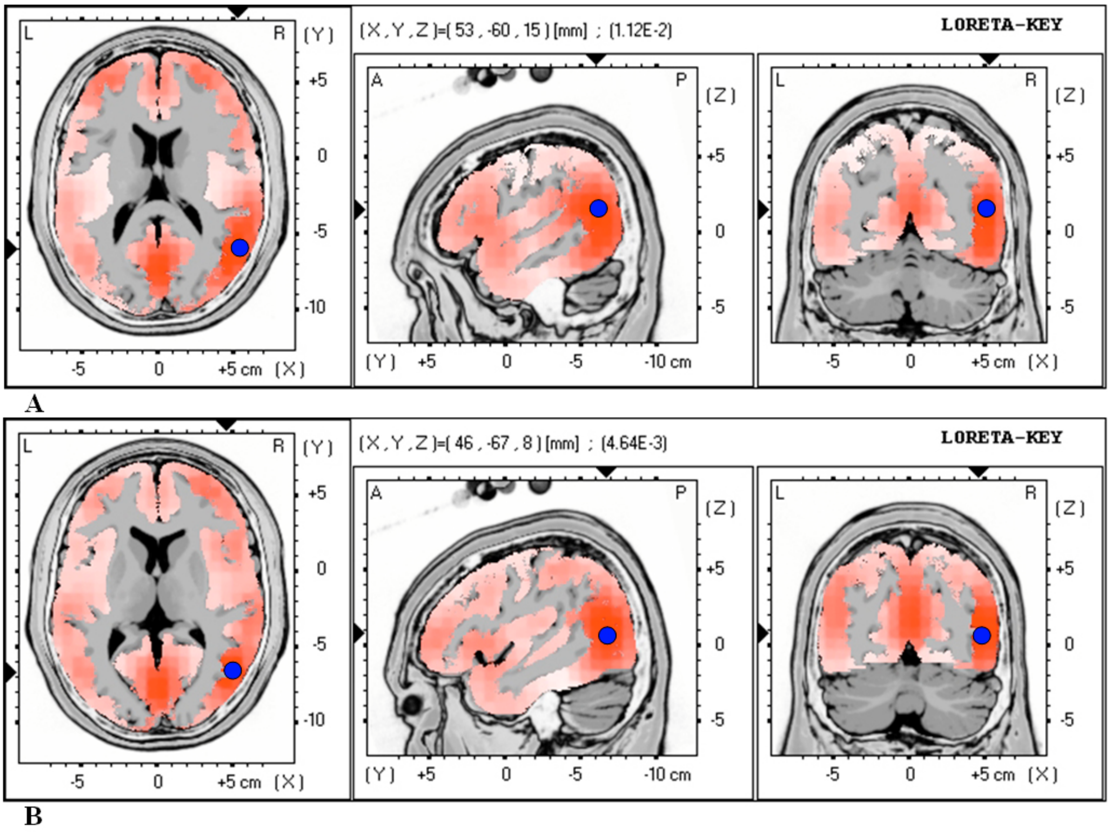

Figure 26: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the rising-to-falling (A) and the falling-to-rising (B) changes of tones with long vowel duration of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

4.2.3. Experiment 6: Preattentive detection of vowel duration change with level-to-contour tone changes

4.2.3.1. Subjects

Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 23-39 (mean 29.7; nine females) and eleven NonS, aged 23-29 (mean 25.8; seven females). The approval of the institutional committee on human research and written consent from each subject were obtained (see appendix I).

4.2.3.2. Stimuli and procedure

Stimuli consisted of four pairs of monosyllabic, Thai words. Speech stimuli were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Three different stimuli were synthetically generated: (1) /k^{h}aam/ - long vowel, level tone; (2) /k^{h}am/ - short, falling; (2) /k^{h}am/ - short, rising. The vowel-duration difference between (1) and (2) was 128 ms (628 vs. 500 ms) and between (1) and (3) was 89 ms (628 vs. 539 ms) with the same intensity used in each stimulus. Five NS listened to the synthesized words and evaluated them all as natural sounding.

The standard (S)/deviant (D) pairs for each experiment, which was randomized across subjects, were shown [Experiment 1] S-1, D-2, [Experiment 2] S-2, D-1, [Experiment 3] S-1, D-3, [Experiment 4] S-3, D-1. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

4.2.3.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.2.3.4. Results

The grand-averaged ERPs show that both level-to-contour and contour-to-level tone changes perception elicited MMN between 196-220 ms with reference to the standard-stimulus ERPs. An ANOVA comparing MMN amplitudes of the S and D yield a main effect of conditions in Experiments 1 ($F_{(3,40)} = 10.58$, $P < 0.0001$) and 3 ($F_{(3,40)} = 13.70$, $P < 0.0001$). In Experiment 2 and 4, however, the S-D differences were not significant (e.g., $F_{(3,40)} = 1.28$, $P = 0.2258$ in Experiment 2, n.s.; $F_{(3,40)} = 0.19$, $P = 0.8559$ in Experiment 4, n.s., for the main effect of conditions). The result showed that long-to-short duration and level-to-falling/rising tone changes elicited a strong MMN bilaterally for NS and NonS, unlike short-to-long duration and falling/rising-to-level tone changes (see Table 11). Furthermore, an across-experiment ANOVA demonstrated no interaction and main effects. The significant difference in MMN amplitudes was not observed between groups across experiments ($F_{(7,80)} = 0.63$, $P = 0.6018$).

Table 11: Mean amplitude (μV)±S.D. of MMN elicited by a vowel duration changes with level-to-contour tones perception in NS and NonS.

Vowel duration change	Native (NS)	Nonnative (NonS)
Long (mid)-to-Short (falling) changes	-1.28±0.40	-0.69±0.36
Short (falling)-to-Long (mid) changes	-1.30±0.20	-0.50±0.21
Long (mid)-to-Short (rising) changes	-0.43±0.11	-0.35±0.50
Short (rising)-to-Long (mid) changes	-0.24±0.19	-0.03±0.21

Source localization analyses were performed using LORETA-Key (Pascual-Marqui, 1994). Table 12 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 212-224 ms. A single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere in Experiment 2 and 3 for NP group and in Experiment 1 and 3 for NS group. Sources were obtained in the MTG of the LH and in the superior temporal gyrus (STG) of the RH for NP group in Experiment 1 and 4 and for NonS group in Experiment 2 and 4, respectively (see Table 12 and Fig. 27- Fig. 34).

Table 12: Stereotaxic coordinates of activation foci during the vowel duration changes with level-to-contour tone perception.

Vowel duration change		Coordinates (mm)			*t* values
	BA	*x*	*y*	*z*	
Native speaker of Thai (NS)					
Long (mid)-to-short (falling) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	3.03
R superior temporal gyrus (STG)	22	53	-60	15	2.11
Short (falling)-to-long (mid) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	2.82
R middle temporal gyrus (MTG)	37	46	-67	8	2.36
Long (mid)-to-short (rising) change					
L middle temporal gyrus (MTG)	21	-52	3	-20	1.75
R middle temporal gyrus (MTG)	37	46	-67	8	1.65
Short (rising)-to-long (mid) change					
L middle temporal gyrus (MTG)	21	-59	-32	1	1.69
R superior temporal gyrus (STG)	22	53	-60	15	1.64
Nonnative speaker of Thai (NonS)					
Long (mid)-to-short (falling) change					
L middle temporal gyrus (MTG)	39	-59	-67	8	2.33
R middle temporal gyrus (MTG)	22	53	-60	15	1.21
Short (falling)-to-long (mid) change					
L middle temporal gyrus (MTG)	21	-59	-32	1	1.40
R superior temporal gyrus (STG)	39	53	-53	15	1.32
Long (mid)-to-short (rising) change					
L middle temporal gyrus (MTG)	21	-52	3	-20	2.95
R middle temporal gyrus (MTG)	37	46	-67	8	2.34
Short (rising)-to-long (mid) change					
L middle temporal gyrus (MTG)	39	-52	-67	8	1.93
R superior temporal gyrus (STG)	22	53	-60	15	1.06

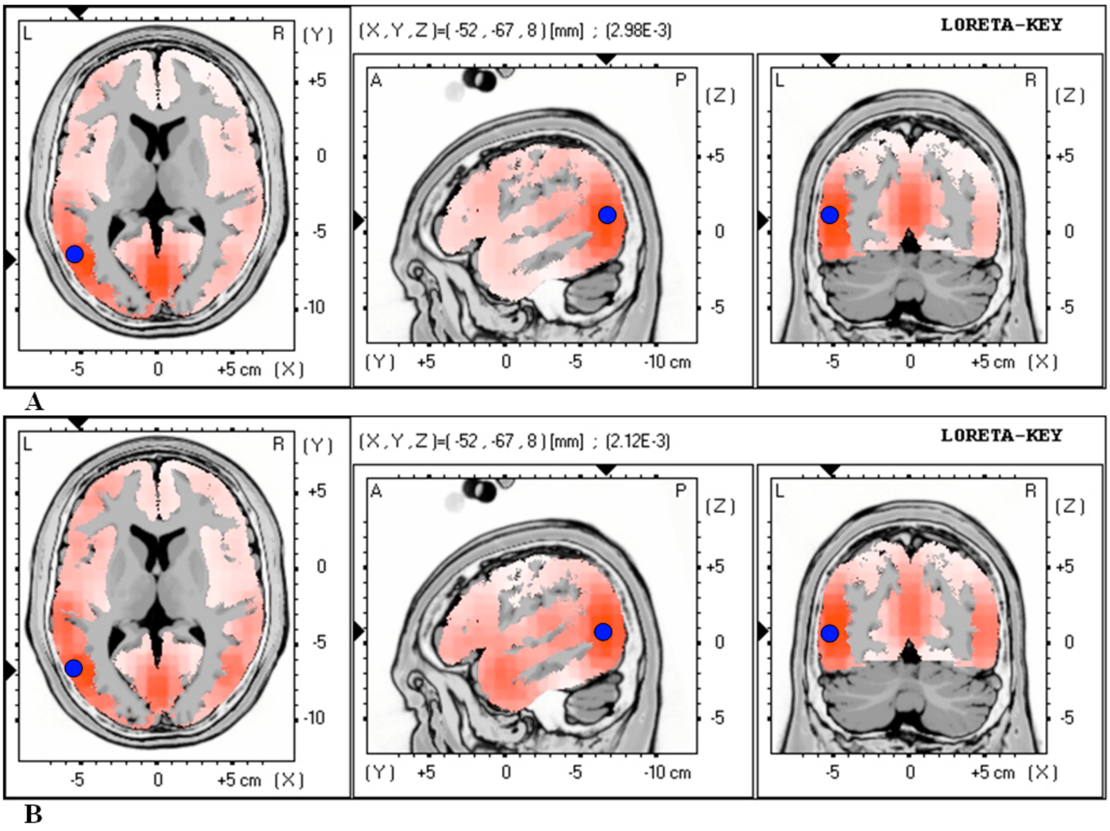

Figure 27: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short changes of vowel duration with level-to-falling tone (A) and the short-to-long (B) changes of vowel duration with falling-to-level tone changes of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

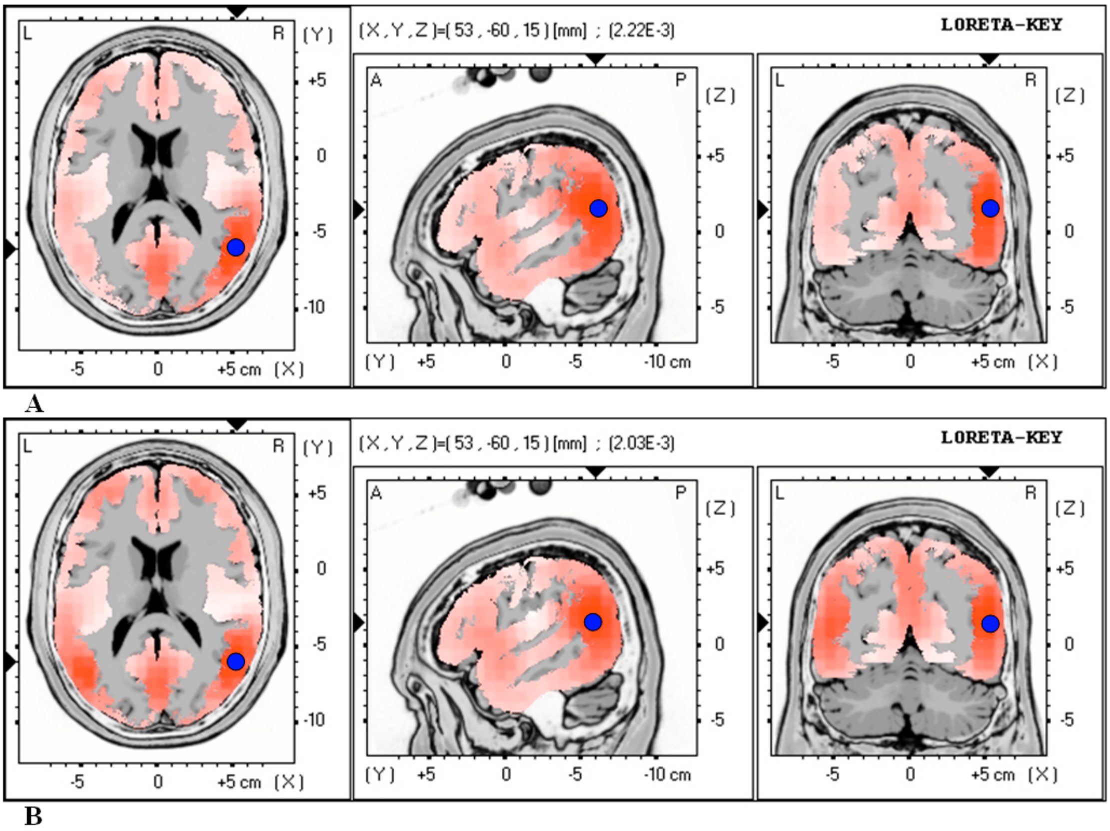

Figure 28: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short vowel duration changes with level-to-falling tone (A) and the short-to-long vowel duration changes with falling-to-level tone (B) changes of vowel duration with level-to-falling tone changes of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

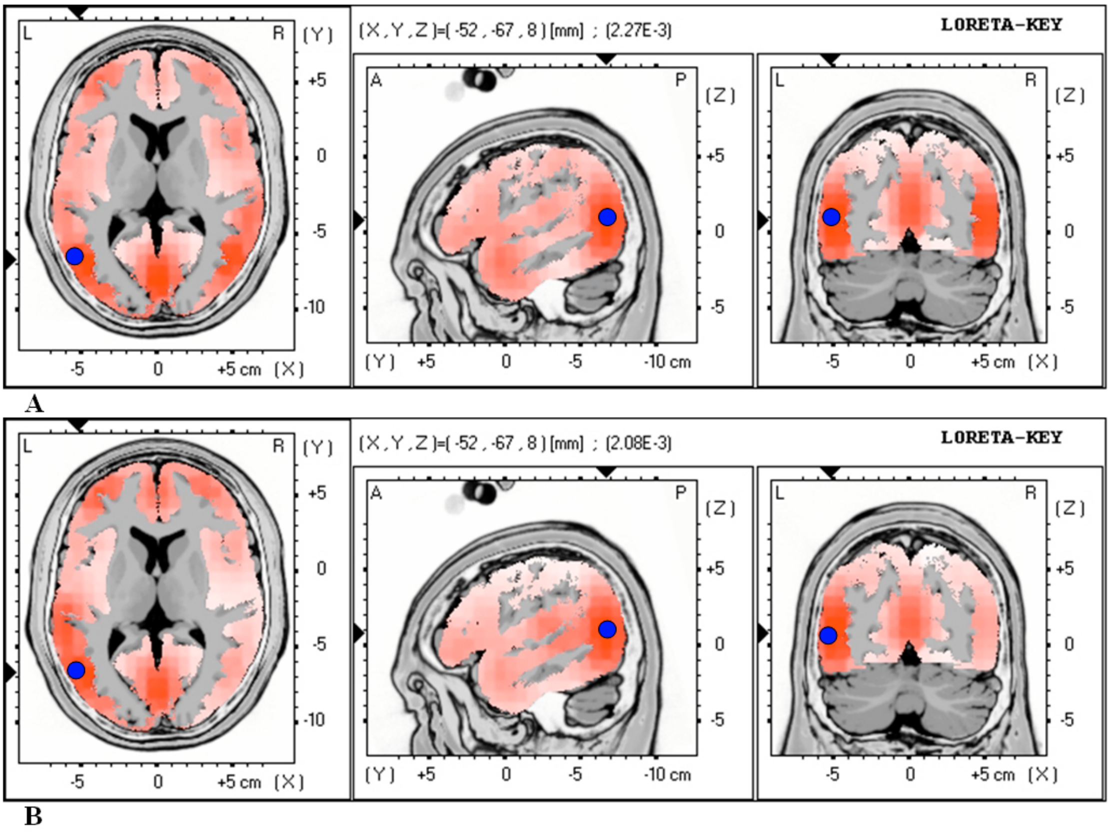

Figure 29: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short vowel duration changes with level-to-rising tone (A) and the short-to-long vowel duration changes with rising-to-level tone (B) changes of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

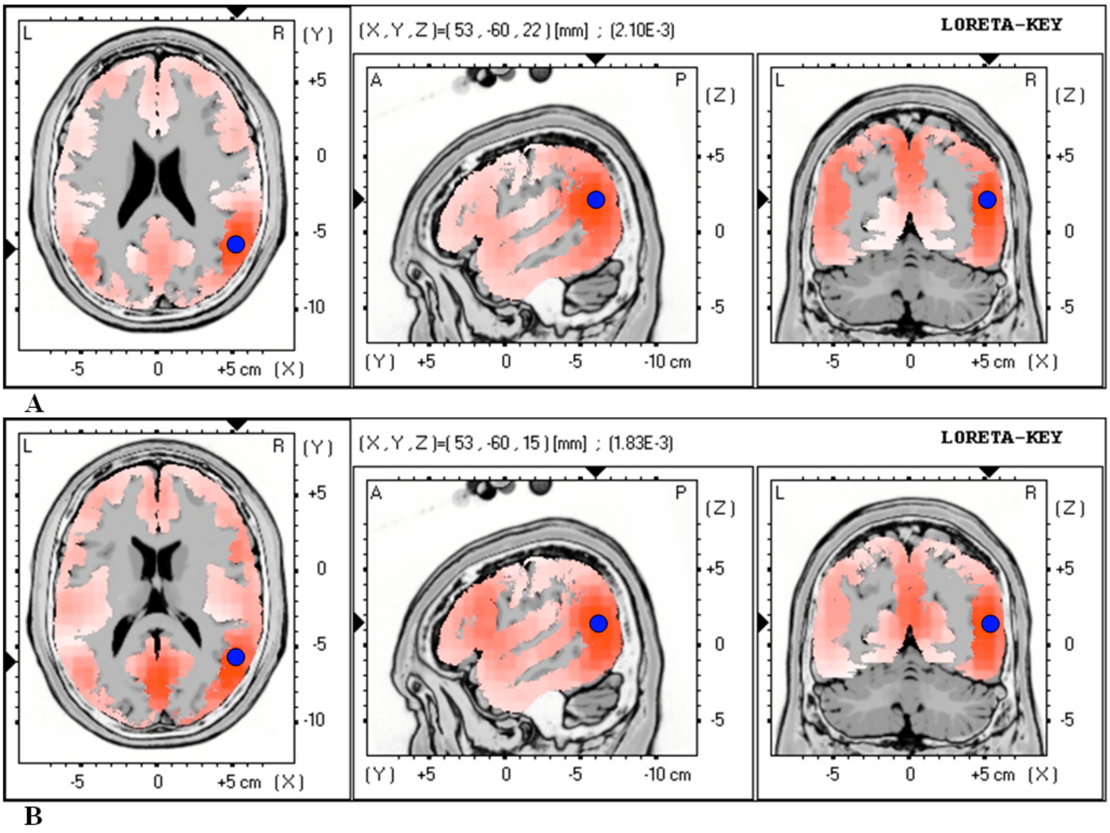

Figure 30: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short vowel duration changes with level-to-rising tone (A) and the short-to-long vowel duration changes with rising-to-level tone (B) changes of NS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

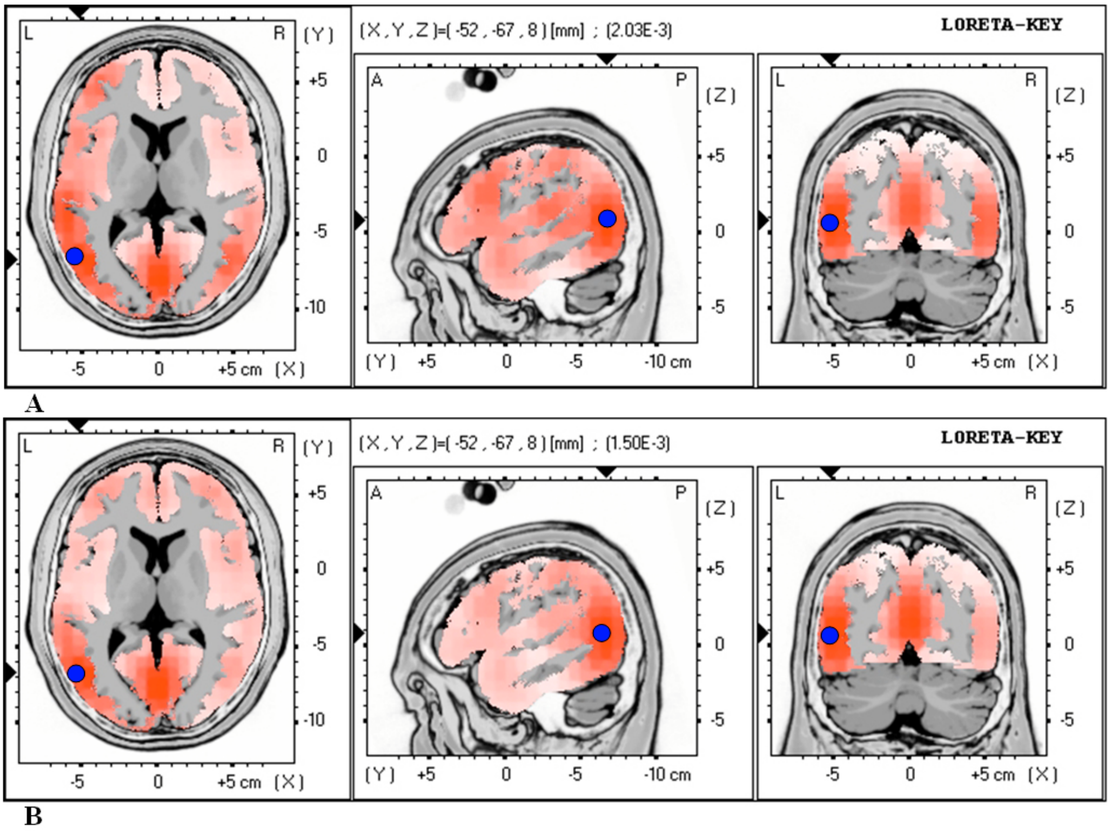

Figure 31: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short vowel duration changes with level-to-falling tone (A) and the short-to-long vowel duration changes with falling-to-level tone (B) changes of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

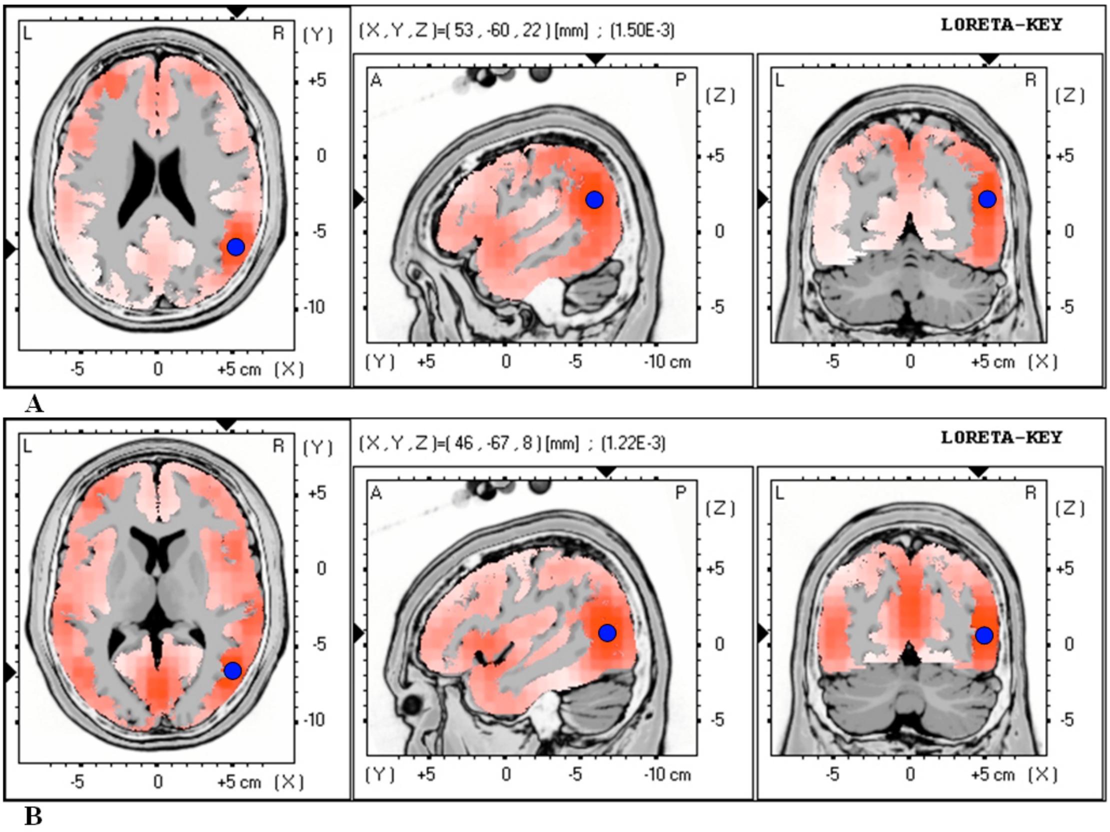

Figure 32: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short vowel duration changes with level-to-falling tone (A) and the short-to-long vowel duration changes with falling-to-level tone (B) changes of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

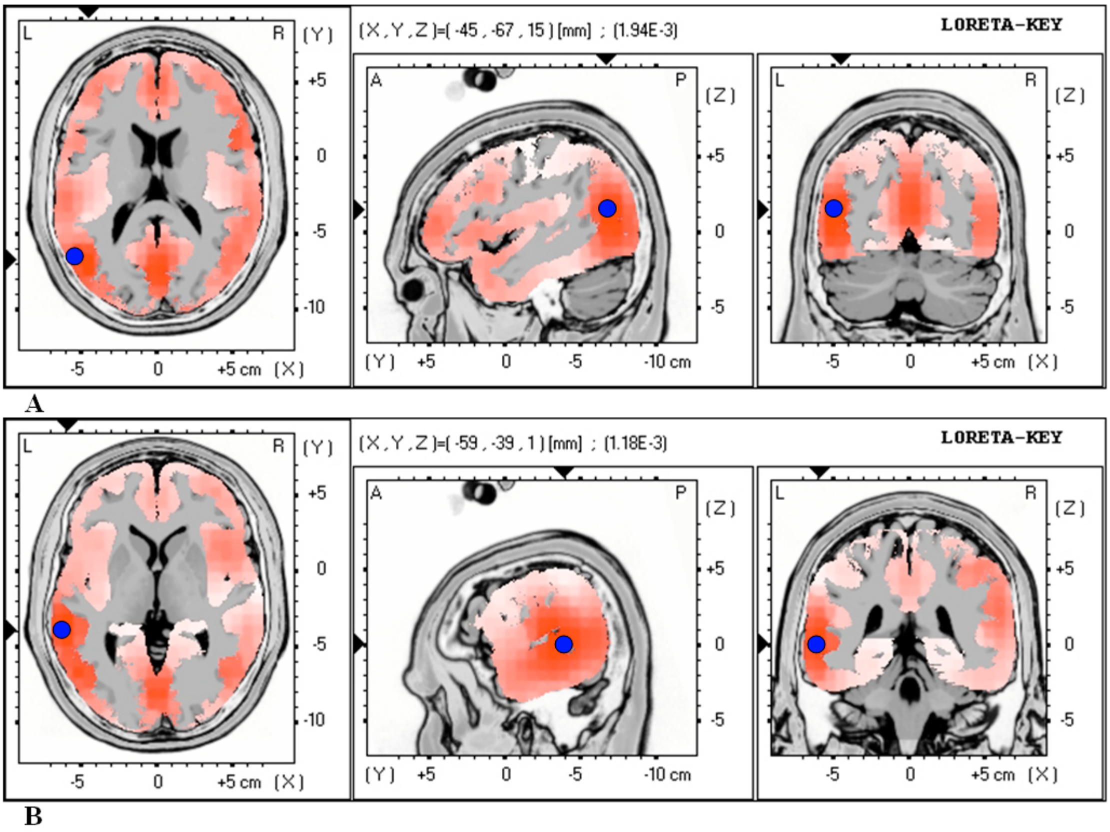

Figure 33: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short vowel duration changes with level-to-rising tone (A) and the short-to-long vowel duration changes with rising-to-level tone (B) changes of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

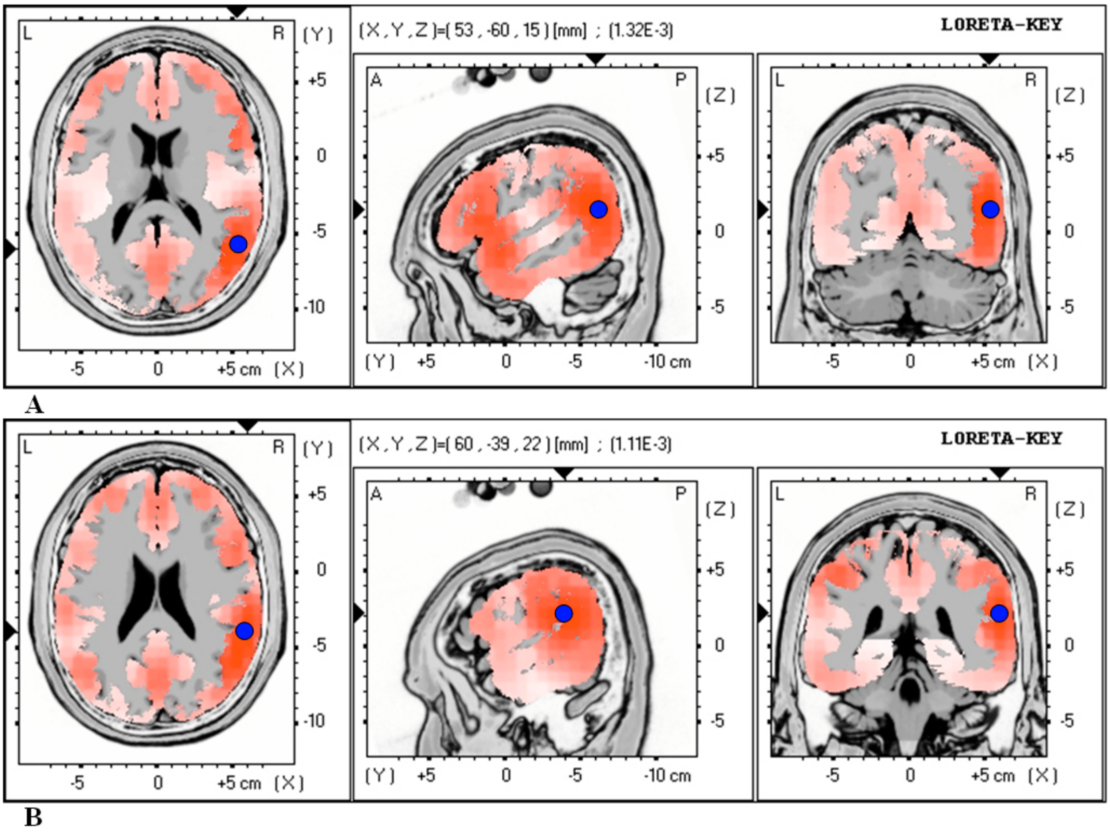

Figure 34: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the long-to-short vowel duration changes with level-to-rising tone (A) and the short-to-long vowel duration changes with rising-to-level tone (B) changes of NonS activated in RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

4.3. Preattentive processing and lateralization of cluster initial consonant

4.3.1. Experiment 7: Preattentive detection of cluster phonemes

4.3.1.1. Subjects

Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 23-39 (mean 27.8; nine females) and eleven NonS, aged 23-29 (mean 27.8; nine females). The approval of the institutional committee on human research and written consent from each subject were obtained (see appendix I).

4.3.1.2. Stimuli and procedure

Stimuli consisted of two pairs of phonemes. They were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50Ω) in dBA using a Brü el and Kjaer 2230 sound-level meter. Two different stimuli were synthetically generated: (1) /k-/ - noncluster phoneme; (2) /kl-/ - cluster. The sound-duration difference between (1) and (2) was 28 ms (62 vs. 90 ms) with the same level pitch and intensity used in each stimulus.

The standard (S)/deviant (D) pairs for each experiment, which was randomized across subjects, were shown [Experiment 1] S-1, D-2, [Experiment 2] S-2, D-1. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

4.3.1.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.3.1.4. Results

The grand-averaged ERPs show that both cluster-and noncluster phoneme changes perception elicited MMN between 140-180 ms with reference to the standard-stimulus ERPs. MMN amplitudes of the S and D was significantly generated a main effect of conditions, as indicated by the t-test comparison between the mean amplitude of MMN at Fz in Experiments 1 ($t_{(10)} = 73.00$, $P < 0.0001$). In Experiment 2, the S-D differences were also significant ($t_{(10)} = 4.71$, $P = 0.0008$, for the main effect of conditions). The result showed that both cluster- and noncluster phoneme changes elicited a strong MMN bilaterally for NS and NonS (see Table 13) Furthermore, an across-experiment ANOVA demonstrated an interaction and main effects. The significant difference in MMN amplitudes was observed between groups across experiments ($F_{(3,30)} = 35.47$, $P < 0.0001$).

Table 13: Mean amplitude (µV)±S.D. of MMN elicited by a cluster and noncluster phonemes changes in NS and NonS.

Cluster and Noncluster Phonemes	**Native (NS)**	**Non-native (NonS)**
Noncluster-to-cluster changes	-1.92±0.07	-0.84±0.34
Cluster-to-noncluster changes	-1.02±0.10	-0.59±0.58

Source localization analyses were performed using LORETA-Key (Pascual-Marqui, 1994). Table 14 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 140-180 ms. In Experiment 1, a single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere for both groups. In Experiment 2, a single source was also obtained in the superior temporal gyrus (STG) of the RH for both groups. No hemispheric difference was discovered in this study (see Table 13 and Fig. 35-36).

Table 14: Stereotaxic coordinates of activation foci during the cluster and noncluster phoneme changes perception.

Consonant phoneme		**Coordinates (mm)**			***t* values**
	BA	*x*	*y*	*z*	
Native speaker of Thai (NS)					
Noncluster-to-cluster change					
L middle temporal gyrus (MTG)	21	-59	-32	1	1.81
Cluster-to-noncluster change					
L superior temporal gyrus (STG)	22	-59	-39	8	1.03
Nonnative speaker of Thai (NonS)					
Noncluster-to-cluster change					
L middle temporal gyrus (MTG)	39	-52	-67	8	2.19
Cluster-to-noncluster change					
L superior temporal gyrus (STG)	42	-59	-32	8	1.14

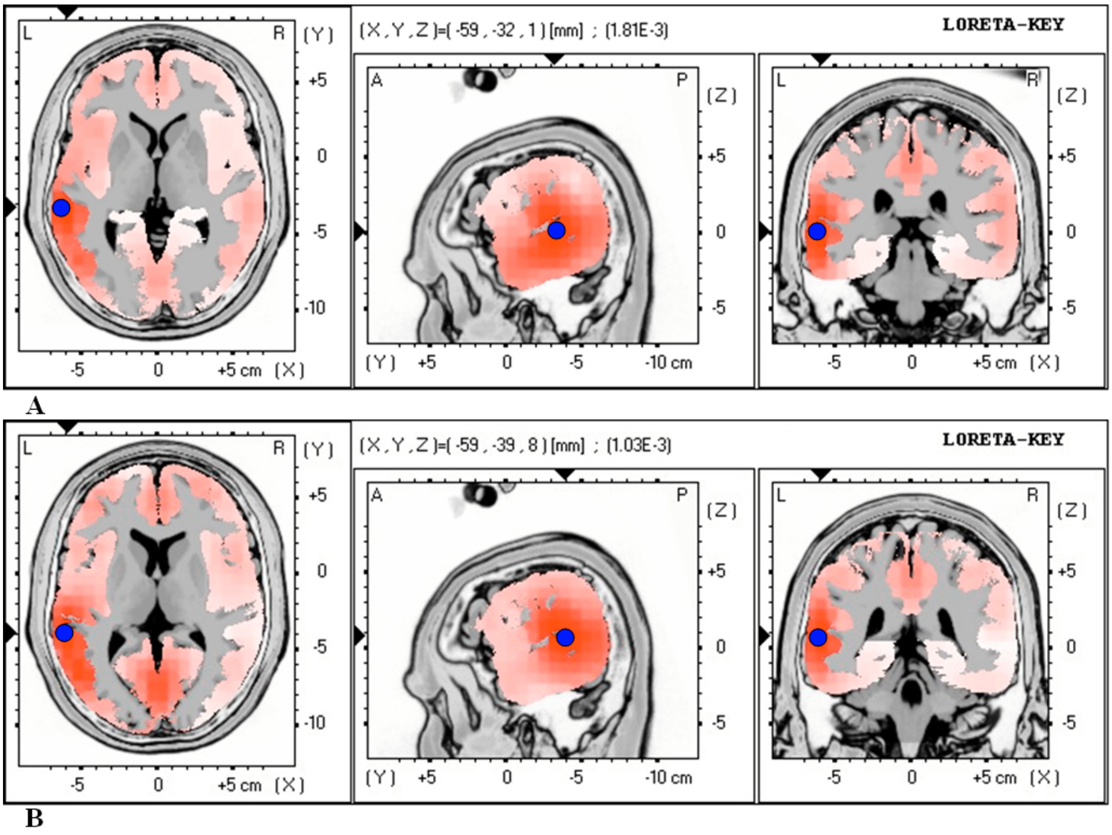

Figure 35: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the cluster (A) and noncluster (B) changes of consonant phonemes of NS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

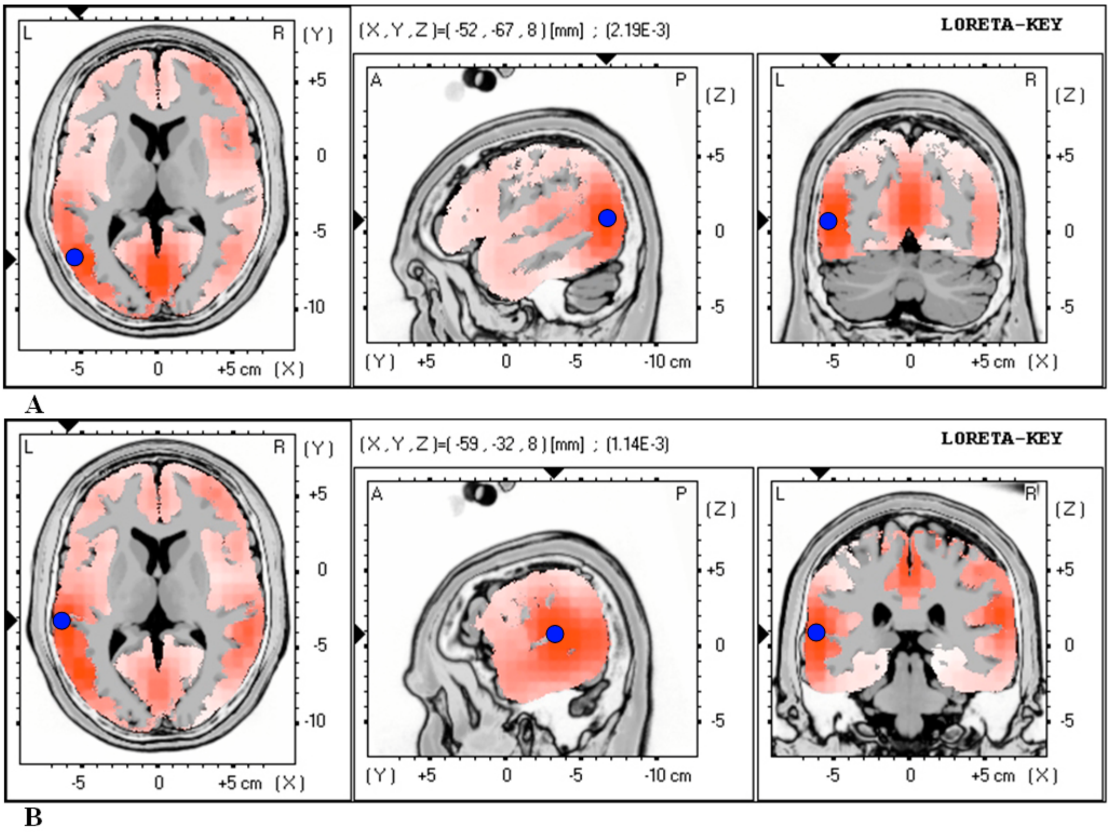

Figure 36: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the cluster (A) and noncluster (B) changes of consonant phonemes of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

4.3.2. Experiment 8: Preattentive detection of cluster initial consonant in monosyllabic Thai words

4.3.2.1. Subjects

Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven NS, aged 23-39 (mean 27.8; nine females) and eleven NonS, aged 23-29 (mean 27.8; nine females). The approval of the institutional committee on human research and written consent from each subject were obtained (see appendix I).

4.3.2.2. Stimuli and procedure

Stimuli consisted of two pairs of monosyllabic, Thai words. Speech stimuli were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Pairs were designed to have similar mid tone. Two different stimuli were synthetically generated: (1) /kang/ - noncluster initial consonant; (2) /klang/ - cluster initial consonant. The sound-duration difference between (1) and (2) was 28 ms (711 vs. 739 ms) with the same level pitch and intensity used in each stimulus. Five NS listened to the synthesized words and evaluated them all as natural sounding.

The standard (S)/deviant (D) pairs for each experiment, which was randomized across subjects, were shown [Experiment 1] S-1, D-2, [Experiment 2] S-2, D-1. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 second (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones, but rather to concentrate on a self-selected silent, subtitled movie.

4.3.2.3. Data analysis

During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample *t*-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

4.3.2.4. Results

The grand-averaged ERPs show that both cluster-and noncluster phoneme changes perception elicited MMN between 208-260 ms with reference to the standard-stimulus ERPs. MMN amplitudes of the S and D was significantly generated a main effect of conditions, as indicated by the t-test comparison between the mean amplitude of MMN at Fz in Experiments 1 ($t_{(10)} = 20.95$, $P < 0.0001$). In Experiment 2, the S-D differences were also significant ($t_{(10)} = 3.15$, $P = 0.0103$, for the main effect of conditions). The result showed that both cluster-and noncluster consonant changes elicited a strong MMN bilaterally for NS and NonS (see Table 15). Furthermore, an across-experiment ANOVA demonstrated an interaction and main effects. The significant difference in MMN amplitudes was observed between groups across experiments ($F_{(3,30)} = 47.02$, $P < 0.0001$).

Table 15: Mean amplitude (µV)±S.D. of MMN elicited by a cluster and noncluster consonants in NS and NonS.

Initial consonants changes	Native	Non-native
Noncluster-to-cluster consonants changes	-2.45±0.66	-2.36±0.45
Cluster-to-noncluster consonants changes	-1.84±1.19	-1.02±0.47

Source localization analyses were performed using LORETA-Key (Pascual-Marqui, 1994). Table 16 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 208-260 ms. In Experiment 1, a single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere for both groups. In Experiment 2, a single source was also obtained in the superior temporal gyrus (STG) of the RH for NS group and the MTG for NonS group. The hemispheric difference was also discovered in this study (see Table 15 and Fig. 37-38).

Table 16: Stereotaxic coordinates of activation foci during the cluster and noncluster initial consonants changes perception.

Consonant phonemes		**Coordinates (mm)**			***t* values**
	BA	*x*	*y*	*z*	
Native speaker of Thai (NS)					
Noncluster-to-cluster consonant change					
L middle temporal gyrus (MTG)	39	-52	-67	8	2.55
Cluster-to-noncluster consonant change					
L superior temporal gyrus (STG)	22	-59	-60	15	1.27
Nonnative speaker of Thai (NonS)					
Noncluster-to-cluster consonant change					
L middle temporal gyrus (MTG)	39	-52	-67	8	7.25
Cluster-to-noncluster change					
L middle temporal gyrus (MTG)	39	-52	-67	8	7.07

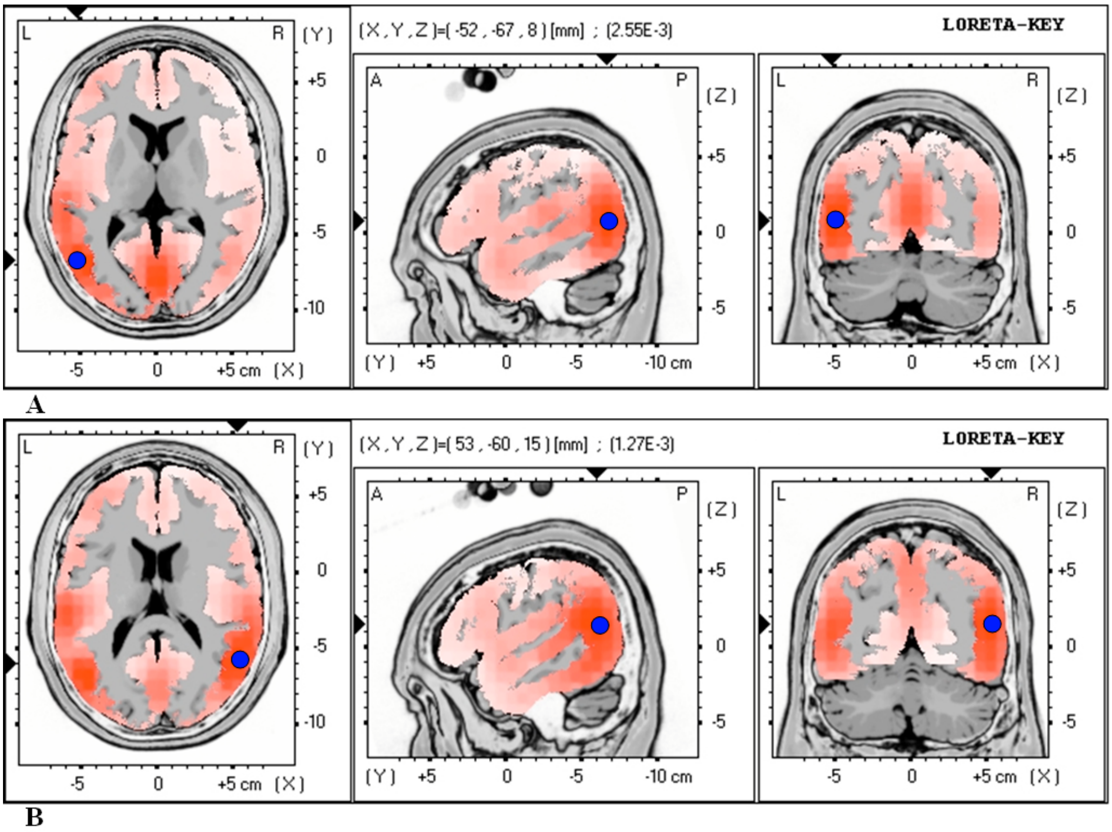

Figure 37: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the cluster (A) and noncluster (B) changes of consonant in CV syllable of NS activated in LH and RH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

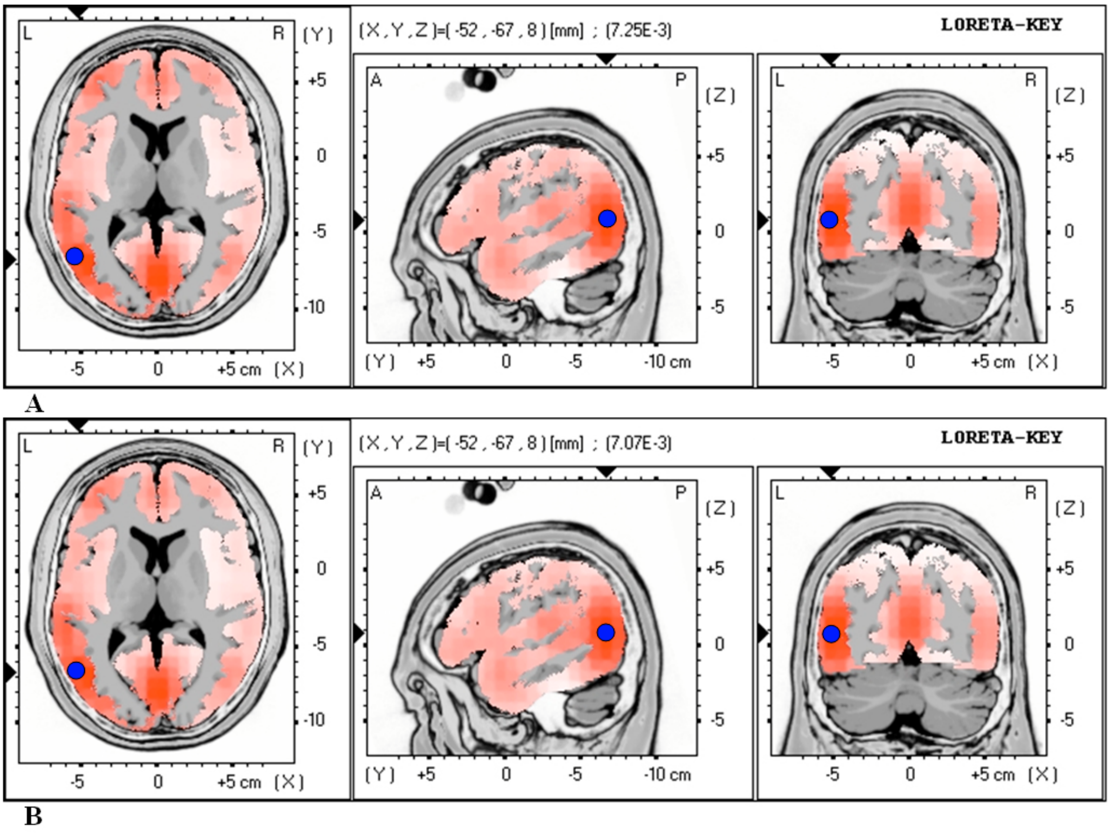

Figure 38: Graphical representation of the low-resolution electromagnetic tomography (LORETA *t*-statistic comparing the event-related potentials for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for the cluster (A) and noncluster (B) changes of consonant in CV syllable of NonS activated in LH. Red color indicates local maxima of increased electrical activity for across- and within-category of vowel change responses in an axial, a saggital and a coronal slice through the reference brain. Blue dots mark the center of significantly increased electric activity.

CHAPTER V
DISCUSSION AND CONCLUSION

In this chapter, the discussions are given in the context of electroneurophysiological studies on preattentive processing.

5.1. Summary of the results

5.1.1. The main findings of the present study show that both across- and within-category vowels changes elicited MMN between 100-140 ms with reference to the standard-stimulus ERPs. The MMN amplitude was statistically significant for both across- and within-category vowel changes. The xyz-values in Talairach space as calculated with LORETA show that the across-category vowel change activated more strongly the left STG whereas the within-category vowel change activated more strongly the right STG.

5.1.2. The duration and tone changes studies show that both long-to-short and short-to-long duration changes with level tone elicited MMN between 184-208 ms with reference to the standard-stimulus ERPs. The long-to-short duration changes with level tone elicited a strong MMN bilaterally for both native and nonnative speaker of Thai, unlike short-to-long duration change with level tone. The xyz-values in Talairach space also show that a source of long-to-short duration changes was estimated to be located in the MTG of each hemisphere for both groups.

5.1.3. The long-to-short duration with falling and rising tone changes perception elicited MMN between 172-264 ms with reference to the standard-stimulus ERPs. The long-to-short duration with falling and rising tone changes elicited a strong MMN bilaterally for both native and nonnative speakers of Thai, unlike short-to-long duration with falling and rising tone changes. Source localization analyses performed using LORETA-Key demonstrates sources were also obtained in the MTG of the LH and in the STG of the RH for both groups.

5.1.4. Both level-to-contour and contour-to-level tone changes perception elicited MMN between 196-220 ms with reference to the standard-stimulus ERPs. The level-to-falling and level-to-rising tone changes elicited a strong MMN bilaterally for both native and nonnative speakers of Thai, unlike falling-to-level and rising-to-level tone changes. Source localization analyses performed using LORETA-Key demonstrates that sources were also obtained in the MTG of the LH and in the superior temporal gyrus (STG) of the RH for both groups.

5.1.5. Both rising-to-falling and falling-to-rising tone changes perception elicited MMN between 212-244 ms with reference to the standard-stimulus ERPs. The rising-to-falling and falling-to-rising tone changes elicited a strong MMN bilaterally for native and nonnative speakers of Thai. Source localization analyses performed using LORETA-Key demonstrates that sources were obtained in the MTG of the LH and in the superior temporal gyrus (STG) of the RH for both groups.

5.1.6. The long-to-short duration change with level-to-falling/rising and falling/rising-to-level tone changes perception elicited MMN between 196-220 ms with reference to the standard-stimulus ERPs. The long-to-short duration and level-to-falling/rising tone changes elicited a strong MMN bilaterally for both native and nonnative speakers of Thai, unlike short-to-long duration and falling/rising-to-level tone changes. Source localization analyses performed using LORETA-Key demonstrates that sources were obtained in the MTG of the LH and in the superior temporal gyrus (STG) of the RH for both groups.

5.1.7. Both cluster-and noncluster stimuli changes perception elicited MMN between 140-180 ms and 208-260 ms for phonemes and CVC syllables, respectively, with reference to the standard-stimulus ERPs. The result showed that both cluster- and noncluster stimuli changes elicited a strong MMN left hemispheric dominant for both native and nonnative speakers of Thai. Source localization analyses performed using LORETA-Key demonstrates that a single source was estimated to be located in the middle temporal gyrus (MTG) of each hemisphere for both groups

5.2. Human auditory evoked mismatch negativity amplitudes vary as a function of category and duration of vowel

The main finding of our study indicates that the prominent response to across- and within-category change of vowel in consonant-vowel syllable elicited MMN peaking at 100-140 ms from stimulus onset. The magnitude of the acoustic difference between the stimulus pairs was reflected by the MMN amplitude. The MMN amplitude was larger in across- than within-category vowel in consonant-vowel syllable contrasts.

Furthermore, results of a tomographic source analysis show a significant bilateral hemispheric activity correlation with the different types of vowel change perception. The LORETA-MMN generator of across-category vowel change was predominantly active in the left auditory cortex while the within-category change of vowel was predominantly active in the right auditory cortex. The left-hemispheric distribution of the source location, estimated by the LORETA in this study, reflected the phonetic and acoustic contrast discrimination of the vowel. This can be the evidence for the left-hemispheric phoneme representation in the human auditory cortex. The converging evidence from factorial statistics of hemispheric effects on LORETA values around the MMN peaks provide further support for the significance of these preattentively activated hemispheric differences in vowel category discrimination.

For a long time the EEG inverse problem could not find a satisfactory solution. As verified by point spread function simulations, sources could be reconstructed at best with some non-systematic localization error. LORETA was found to have zero localization error (Pascual-Marqui et al., 1994), hence it finally solved exactly the EEG inverse solution problem (as far as localization error in the absence of noise is concerned). The present study thus used standard LORETA space consisting in 2394 voxels of size 7x7x7 mm. The spatial dispersion of the reconstruction with 21 electrodes is considerable with the zero localization. With respect to the usefulness of a new tomographic source analysis, the present data showed that the LORETA can locate multiple non-dipolar sources accurately to demonstrate the dynamic, alternating between left and right homologous regions for different category of vowels in language processing.

Moreover, using the current density measures, we demonstrate that different stages of vowel category processing in consonant-vowel syllable may be associated with left or right hemisphere prominence, in addition to bilateral activation in superior temporal gyrus. Left-sided activation occurs predominantly to the across-category vowel change, while its right-sided activation occurs mainly for the within-category vowel change. This suggests that differential superior temporal gyrus activation is due to phonetic rather than lexical-semantic processes. Bilateral activation in superior temporal gyrus might be involved in encoding combinations of temporal and spectral features as phonetic representations. The across-category change of vowel was specific to changes in spectral cues, but the within-category of vowel was specific to changes in temporal cues. Early stages, associated with across-category change of vowel perception, thus, were demonstrated by the left hemisphere. This was followed by activation of the homologous right hemisphere region for within-category vowel change (Sinai and Pratt, 2003; Coney and Evans, 2000; Knosche et al., 1999). The different pattern of activation to within-category and across-category changes of vowels support the view that information transfer in the right and left hemispheres are for analysis of different aspects: physical features within the same phoneme boundary in the right and independent phonemes in the left hemisphere.

Furthermore, when the vowels of the stimulus pair belonged to different categories, in other word, were in the across-category condition: /O/-/i/ as in /pO/-pi/), the assignment of the stimulus as standard or deviant could have created a contrast effect. The assignment, on the other hand, played no role in the same or within-category one (i.e. /O/-/o/ as in /pO/-/po/). This might have resulted from the fact that the phonetic status was not obvious to the present subjects because the area of the /pO/ as stimulus standard is near the /po/ as stimulus deviant. The results thus showed that the MMN depended primarily on the acoustic differences between the members within the vowel pairs. In other words, the MMN amplitude was larger for a vowel pair with a large acoustic distance as compared with a pair with a smaller distance in the acoustic space.

The present results support a model suggesting that in the intact brain, each hemisphere relies on earlier processing stages in both hemispheres and on the ongoing evolving stimulus, which contribute to searching the lexicon. This model is compatible with separate right and left hemispheric distributions of neural networks, interconnected and using the processing of the other hemisphere's output as input for its computations. The time scale of inter-hemispheric transfers suggests that information is transferred back and forth between left and right homologous regions. The extensive information exchanges between right and left homologous sites suggest that as a word is evolving over time, right and left hemisphere regions contribute and interact to create efficient linguistic parallel processing (Sinai and Pratt, 2003). This may include the transfer of information to the right hemisphere to free resources in the left one to focus on the meanings (Sinai and Pratt, 2003; Coney and Evans, 2000) or to process further input as the word's sounds progress over time (Knosche et al., 1999). Thus, the bilateral activation in STG of different category vowel change perception was compatible with previous study (Jäncke et al., 2002) that two possible time courses may be involved: (1) both hemispheres are active in parallel with the left hemisphere dominant for different phonemes boundaries and (2) hemispheric dominance may alternate to play a specific role at different stages of linguistic processing. This supports the hypothesis that the part of superior temporal gyrus contains neural networks specialized for language processing (Jäncke et al., 2002).

Our results are supported by the MEG studies showing that the processing of a longer vowel (600ms) was mainly lateralized on the left hemisphere (Eulitz et al., 1995; Obleser et al., 2001). However, contradicting evidence was also found in previous reports which employed speech sounds with shorter duration: the left hemispheric predominant MMN was not systematically obtained (Eulitz et al., 1995; Tervaniemi et al., 1999; Aaltonen et al., 1994; Vihla and Salmelin, 2003). It has been proposed that isolated semisynthetic vowels with short duration in a repetitive manner are not processed fully as phonemes in the subject's brain (Kasai et al., 2001). Another possible reason for the discrepancy between our study and previous studies is the naturalness of stimuli. It has been hypothesized that the categorical perception of vowels is increased by the complexity of the synthesis and thus affected by the listener's discrimination behavior (Savela et al., 2003). By the use of natural speech of vowel in consonant-vowel syllable, the present data show that the speech-sound naturalness already affects the earlier e.g., preattentive level of speech perception.

5.3. Auditory cortex responses to vowel duration and pitch changes

The Main finding is that for both groups, vowel-duration shortening and falling tone elicited prominent MMN components, whereas vowel-duration lengthening and level tone. The same results were obtained in MEG studies using both tone and Japanese words (Inouchi et al., 2003, 2004, 2005). The present findings are in accord with previous vowel experiment that showed that shortening elicited a larger MMN than lengthening (Jaramillo et al., 1999; Inouchi et al., 2003). This suggests that at the point of stimulus disparity or thereafter, change detection of vowel lengthening is somehow compromised. It is reasonable to speculate that the continued auditory processing required for the longer vowel interferes with or masks the detection mechanism underlying the MMN (Inouchi et al., 2003). In contrast, the current findings contradict previous tone studies that reported a clear MMN elicited by both duration increments and decrements (Näätänen et al., 1989) and a larger MMN elicited by increments than decrements (Jaramillo et al., 1999). However, the stimulus and presentation between these studies differed considerably; notably the referred to tone study contained stimuli of short durations compared to our longer "word length' stimuli and this might explain the disparity.

The present results parallel the finding in previous studies (Inouchi et al., 2002, 2003, 2004). The present study, the detection of vowel duration and tone changes was most likely acoustically driven rather than semantically driven, such that the stimuli were processed without any access to semantic information. The acoustic aspect in the absence of phonetic or higher-order properties may account for why NonS had similar neuronal responses to NS subjects. Additionally, the length of our stimuli may have contributed to the lack of difference in MMN between NS and NonS. Temporal integration of acoustic energy is therefore purposed to take place in the time range of the initial 200 ms of the stimulus (Michie, 2001; Näätänen, 1992). After such time, auditory processes have integrated the sensory components as a unitary event in sensory memory (Näätänen, 2001). Nenonen et al., (2003) who found cross-linguistic differences for processing shortened durations used deviants that were shortened from 200 to 150 ms, a deviation that may have straddled the integration threshold for one or both groups. Therefore, our stimuli (e.g., Thai words) were considerably longer and most likely were integrated in sensory memory. The differences between these studies provide the impetus for future investigations of duration processing and temporal integration differences across language groups.

The MMN component was also found to be more sensitive to tone falling than rising and leveling for both groups. To the extent that both shortening-duration and falling tone conditions reduce acoustic energy in the deviant compared to the standard, similar sensitivity in the MMN is not surprising. The delay in the MMN for falling tone, compared to that of shortening duration, may be due to the incremental spectral deviation, as opposed to the rapid change (omission) in the shortened stimulus (Inouchi et al., 2003).

The source analyses of the MMN components suggest that shortened-duration and tone changes may elicit activity in the superior temporal sulcus (STS) of each hemisphere for both NS and NonS. Source modeling using a single equivalent dipole approach has well-recognized spatial limitations, perhaps accounting for the discrepancy between current findings and previous reports of MMN/MMF generators in the planum temporale. However, while there existed no hemispheric differences in the current study, the present studies are in line with previous f-MRI (Binder et al., 2000) and MEG (Inouchi et al., 2003) reports that found the left posterior superior temporal gyrus to be activated by the pre-attentive detection of acoustic changes in non-speech (tones) and speech (CV syllables). However, a previous study (Inouchi et al., 2003) purposed that it may be that duration changes in tones activated more inferior brain regions centered in STS in contrast to spectral changes centered in the lateral fissure, but such inferences are beyond the scope of the study, since only one subject in each group had MR imaging.

One might expect language-specific effects on the elicitation of the MMN in speech, since Thai is a tonal language and English is a stress-accent language. Although a tendency towards stronger MMN in NonS is observed, the current findings do not support such an expectation in that no statistically significant difference in MMN amplitudes was found between groups at an acoustic level. Previous ERP studies at a phonetic level demonstrated that the MMN was enhanced in Finnish subjects by their first-language (Finnish) phoneme prototype rather than a non-prototype (Estonian) (Nää tänen et al., 1997) and that the MMN for a vowel contrast in Finnish was not generated in native Hungarian speakers with no knowledge of Finnish (Winkler et al., 1999), implying that the MMN reflects language-specific memory traces formed by early and extensive exposure to a first language. However, language-specific word-related MMN/MMF components at acoustic and phonetic levels remain to be investigated in future studies.

This study confirmed previous results showing that MMN was elicited by both stimulus duration increments and decrements (Jaramillo et al., 1999; Joutsiniemi et al., 1998; Näätänen et al., 1989). MMN generally increased as a function of the degree of discrepancy from the standard-stimulus duration. Nonetheless, the fact that the MMN was still elicited by such an extremely brief sound suggests that the MMN generator is indeed sensitive to even very small amounts of stimulus energy. Consequently, the finding that the MMN amplitude increases as a function of discrepancy in duration and tone change from the standard suggests that the MMN is generated by a process which is a neural code of the degree of stimulus change and accord with results from similar paradigms using other features (Näätänen et al., 1989).

5.4. Preattentive processing and lateralization of cluster consonant

In the present study, the MMN was recorded from both NS and NonS in responses to the synthesized stimuli /k/ and /kl/ as phonemes and as in CV syllables. Both experiments evaluated the effect of stimulus duration (long vs. short) on the MMN responses. A clear MMN in the left hemisphere was observed only for the cluster stimuli but not for the noncluster stimuli. The current study did not show the duration effect and the MMN was always elicited independent of stimulus duration. It is thus confirmed that the duration effect on the MMN comes from the backward masking of the consonant by the following double consonant. As it is well established that the MMN amplitude indexes the change detection accuracy (Näätänen, 2001), the larger MMN amplitude to the noncluster-to-cluster stimuli change in the native speakers of Thai suggests more accurate cluster change detection in Thai than in the second-language users. Consistent to this the MMN to noncluster stimuli was significantly smaller than those for luster stimuli. The source analysis for the cluster stimuli revealed that the source for cluster stimuli was located significantly more superior than for noncluster stimuli in the left hemisphere. In addition, the estimated source strength was not difference between cluster and noncluster stimuli. It is thus the source location rather than its signal strength that might be a primary reason for the reduced responses for noncluster with cluster stimuli. These results may be comparable to those of a previous report by Tiitinen et al. (1993) indicating that the MMN sources as well as N100m have a frequency-dependant tonotopy. We propose that the tuned processing of cluster consonant sounds in the Thai and the lack of it in the second-language users is caused by the different roles of sound frequency and duration in the subjects' native language. Hence the fine tuning in the processing of the cluster consonant may be inhibited at the pre-attentive level.

One possible explanation of this hemispheric discrepancy in the MMN cluster effect is that the temporal window of integration (Näätänen and Winkler, 1999; Yabe et al., 1998, 2001; Sussman et al., 1999) differs between two hemispheres. The left hemisphere is suggested to be more sensitive than the right hemisphere to high-frequency temporal patterns of sounds including the stimulus duration. In fact, the perceptual right-ear advantage, regarded as an index of the left hemisphere dominance, can be observed for high-frequency sounds and this advantage is reversed for low-frequency sounds (Deutsh, 1994). This may relate to the lack of frequency and duration effect in the right hemisphere, since the cluster /kl-/ and noncluster /k-/ distinction mainly rests on the difference of the relatively high frequencies for each of the stimuli. Additionally, the sources for the cluster stimuli were located more posterior and superior in the left hemisphere than in the right hemisphere. The results replicated previous studies (Koyama et al., 2000a; Koyama et al., 2003) and other studies (Alho et al., 1998; Rinne et al., 1999; Tervaniemi et al., 1999; Levänen et al., 1996; Koyama et al., 2000b). For the source strength of the cluster and noncluster stimuli, no hemispheric difference was observed. I should be also noted that MMN to cluster stimuli showed significant left hemisphere dominance. These results might be in accordance with the findings of Mathiak et al. (1999) rather than studies reporting the left lateralized MMN for speech sounds (Näätänen et al., 1997; Tervaniemi et al., 1999; Shtyrov et al., 2000; Koyama et al., 2000b).

Zatorre et al. (2001) also reports the MMN from the /l/-/r/ contrast in Japanese subjects. They used synthesized speech sounds and the initial part of their stimuli was extended. The syllables started with 155-ms steady formant structure followed by a 100-m F3 transition with the total stimulus duration being 400 ms. Such an extended consonant part is assumed to be helpful in reducing the masking effect from the following vowel. The masking effect has been regarded as evidence of temporal integration occurring within the time window of 150-200 ms (Näätänen and Winkler, 1999; Yabe et al., 199, 2001; Sussman et al., 1999) corresponding to the extended consonant. On the other hand, studies using a shorter consonant part and a longer vowel part do not report MMN from the /l/-/r/ contrast in Japanese subjects (Phillips et al., 1995; Imaizumi et al., 1999).

In summary, the present study confirmed that the frequency and duration effect of MMN (i.e., present for the cluster stimuli, absent for the noncluster stimuli) for Both NS and NonS subjects using the new subtraction method. Hence it is shown that this effect does not merely come from the difference in the responses to standard sounds. The cluster and noncluster stimuli did not show the duration effect. The MMN was always elicited independent of stimulus duration. The present findings are consistent with other MMN/MMNm studies that manipulated the initial transitional part of the consonant part (Zhang et al., 2001; Bradlow et al., 1999). It should be noted that in order to examine the masking effect on the consonant by the following consonant, the consonant truncation approach may be better than extending the consonant part for two reasons. First, consonant extension can alter the quality of speech sounds and can result in “unnatural” speech sounds. Second, consonant extension changes the amount of acoustical difference between two sounds, whereas the consonant truncation does not. It is generally expected that the larger the difference between the deviant and standard sounds, the larger MMN amplitude. Hence the change in MMN can thus be more specially related to the masking effect using consonant truncation than using consonant extension.

5.5. Concluding remarks

5.5.1. The vowel across-category-change perception in consonant-vowel syllable elicited MMN peaking at 100-140 ms from stimulus onset. As assessed by LORETA, the MMN was predominantly generated in the left auditory cortex, thus emphasizing the role of the left hemisphere in the auditory preattentive processing of vowel across-category-change perception in consonant-vowel syllables.

5.5.2. The MMN component is sensitive to vowel shortening rather than lengthening and to pitch falling rather than rising and leveling. Automatic detection of changes in vowel shortening and pitch falling is a useful index of language universal auditory memory traces.

5.5.3 The prominent MMN component was elicited by cluster consonant changes in each hemisphere for both subject groups. The MMN is more sensitive to cluster consonant than noncluster consonant in natural language processing. Cluster consonant change is considered particularly salient cues for pre-attentive auditory change detection Thai words.

5.5.4. While the right hemisphere is predominant in the perception of the non-native speech sounds, the left hemisphere is predominant in the perception of the native speech sounds. The hemispheric pattern of the cortical activity related to native speech-sound discrimination is different from that involved in non-native speech-sound discrimination processing: the left-hemispheric dominance in speech perception can be observed already at its early stage and may be explained by pre-existing long-term memory traces (the acoustic templates of perceived signals) for the native speech sounds formed in the dominant hemisphere.

5.5.5. When the speech signal is masked by the segmental (e.g., consonant and vowel) phonological unit of lexical tones, the involvement of the left (dominant) hemisphere in discriminating speech signals increases while that of the right hemisphere increases in the suprasegmental (e.g., prosodic) one. The activity-source locations of a number of responses are in the segmental component different from those in the suprasegmental component. So, the lateralization of the speech-sound discrimination function is distributed between the cerebral hemispheres, which presumably involved the recruitment of additional neural structures in the left hemisphere in perception of speech and the right hemisphere in perception of prosodic component of speech.

5.5.6. The MMN presumably reflect the early stage of speech processing in the human brain. The MMN reflects an early, pre-attentive, automatic speech processing. So, from the known early auditory-cortex responses to sounds, only the mismatch negativity seems to be sensitive to the hemispheric lateralization of the speech function. An advantage of the possible application of MMN as a measure of speech lateralization is that it can be used, unlike behavioral measures, with any subject groups, including patients unable to communicate or concentrate on a test task. The MMN might be of potential interest as a technique of evaluating speech-processing lateralization, since its measurement is non-invasive, relatively inexpensive (especially in case of the EEG), and applicable to any subjects or patients. This should be therefore further studied.

5.5.7. The present study has added physiological evidence to earlier psychological and clinico-anatomical evidence that functional characteristics of auditory stimuli differentially influence the brain circuits used at higher cortical stages for speech processing. The present study has also demonstrated that the grand enterprise of mapping language onto the human brain can be vitally enhanced by MMN studies.

BIBLIOGRAPHY

Aaltonen O, Eerola O, Lang AH, Uusipaikka E and Tuomainen J. Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity. *J Acoust Soc Am.* 1994; 96:1489-1493.

Ahonnisk J, Cantell M, Tolvnen, Lyytinen H. Speech perception and brain laterality: the effect of ear advantage on auditory event-related potentials. *Brain lang.* 1993; 45: 127-146.

Alexander JE, Polich J. P300 differences between sinistrals and dextrals. *Cogn. Brain Res.* 1995; 2: 277-282.

Alho K. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear Hear*. 1995; 16: 38-51.

Anastasopoulos G. Linksseitige Hemiplegie mit Alexie, Agraphie, und Aphasie bei einem polyglotten Rechtshander. *Z Neurol.* 1959; 179: 120-144.

Angelergues R, Hecaen H, Djindjian R. Un cas d'aphasie croisee. *Rev Neurol.* 1962; 107: 453-454.

April RS, Han M. Crossed aphasia in a right-handed bilingual Chinese man, a second case. *Archives of Neurology*. 1980; 37: 342-346.

April RS, Tse P-C. Crossed aphasia in a Chinese bilingual dextral. *Archives of Neurology* 1977; 34: 766-770.

Azou P, Eustche F, Etevenon P, Platel H. Topographic EEG activations during timbre and pitch discrimination tasks using musical sounds. *Neuropsychologia.* 1995; 33: 25-37.

Barraquer-Bordas L, Mendilaharsu C, Peres-Serra J. Estudio de dos casos de afasia cruzada en pacientes manidextros. *Acta Neurol Lat Am.* 1963; 9: 140-148.

Barrett SE, Rugg MD. Evoked potential asymmetries related to handedness and side of stimulation. *Neuropsychologia.* 1989; 27: 539-548.

Baudoin-Chial S. Hemispheric lateralization of Modern Standard Chinese tone processing. *Journal of Neurolinguistics*. 1986; 2: 189-199.

Behrens SJ. Characterizing sentence intonation in a right hemisphere-damaged population *Brain and Language*. 1989; 37: 181-200.

Benson DF. *Aphasia, alexia, and agraphia*. New York: Churchill Livingstone; 1979.

Berry WR. A psychometric reconsideration of the Token Test. Paper presented at *the Third Annual Conference on Clinical Aphasiology*, Albuquerque; 1973.

Bolinger D. *Language: The Loaded Weapon*. Longman Group: London; 1980.

Botez ML, Wertheim, N. Expressive aphasia and amusia following right frontal lesion in right-handed man. *Brain* 1959; 82: 186-202.

Bregman AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT, Press; 1990.

Broca P. "Du siege de la faculte du language articule," *Bulletin de la Scociete d' Anthropologie*. 1865; 6: 377-93.

Brown J, Wilson FR. Crossed aphasia in a dextral. *Neurology.* 1973; 23: 907-911.

Bryden MP. Pattern of cerebral organization. *Brain Lang*. 1983; 20: 249-62.

Bryden MP, Ley RG. *Right-hemispheric involvement in the perception and expression of emotion in normal humans*. In K. Heilman and P. Satz (Eds.), Neuropsychology of human emotion, pp. 6-44. New York: The Guilford Press; 1983.

Cancelliere, Kertesiz Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and Cognition*. 1990; 13: 133-147.

Chobor KL, Brown JW. Phoneme and timbre monitoring in left and right cerebrovascular accident patients. *Brain Lang*. 1987; 30: 278-284.

Clarke B, Zangwill O. A case of "crossed aphasia" in a dextral. *Neuropsychologia.* 1965; 3: 81-86.

Collins L, Holmes C, Peters TM, Evans AC. Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping*. 1995; 3 (3): 190-208.

Connolly JF. The influence of stimulus intensity, contralateral masking and handedness on the temporal N1 and the T complex components of the auditory N1 wave. *Electroenceph. Clin. Neurophysiol*. 1993; 86: 58-68.

Cooper WE, Klouda G. Intonation aphasic and right-hemisphere damaged patients. In J. Ryalls (Ed.), *Phonetic approaches to speech production in aphasia and related disorders*, pp. 59-77. Boston: Colleger-Hill; 1987.

Cooper WE, Sorensen. *Fundamental frequency in sentence production.* New York: Springer-Verlag; 1981.

Coney J, Evans KD. Hemispheric asymmetries in the resolution of lexical ambiguity. *Neuropsychology* 2000; **38:**272-282.

Crystal D. Non-segmental phonology in language acquisition: Review of the issues. *Lingua.* 1973; 32: 1-45.

Crystal D. *The English Tone of Voice.* St. Martin's Press: New York; 1975.

Cutler A, Chen H-C. Lexical tone in Cantonese spoken-word processing. *Perception and Psychophysics.* 1997; 59: 165-179.

Cutler A, Clifton C. Comprehending spoken language: A blueprint of the listener. In C. Brown and P. Hagoort (Eds.), *The Neurocognition of Language*, pp. 123-166. New York: Oxford University Press; 1999.

Cutler A, Otake T. Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America*. 1999; 105: 1877-1888.

Danly M, Shapiro B. Speech prosody in Broca's aphasia. *Brain and Language*, 1982; 16: 171-190.

Darwin CJ. *"On the Dynamic use of Prosody in Speech Perception",* In A Cohen and S.G. Nooteboom (Eds.), pp. 178-93. Berlin: Springer-Verlag; 1975.

Dax M. "Lesions de la moitie gauche de l' encephale coincidant avec l'oubli des signes de la pensee. Lu au Congres meridional Tenu a Montpellier en 1836," published by Gustave Dax (1865) *Gazette Hebdomadaire de Medecine et de Chirurgie.* 1836; 33 (227): 62.

Deacon T. *Connections of the inferior periarcuate area in the brain Macaca fascicularis: an experimental and comparative investigation of language circuitry and its evolution.* Ph.D. Thesis, Harvard University; 1984.

Demany L. Auditory stream segregation in infancy. *Infant Behav*. Dev. 1982; 5: 261-276.

DeRenzi E, Vignolo LA. The Token Test: a sensitive test to detect receptive disturbances in aphasics. *Brain* 1962; 85: 665-678.

DeRenzi E, Vignolo LA. *The Token Test (TT) in Evaluation of Appraisal Techniques in Speech and Language Pathology* edited by Darley FL. Sydney: Addison-Wesley Publishing Company; 1979.

Diesch E, Eulitz C, Hampson S and Ross B. The neurotopography of vowels as mirrored by evoked magnetic measurements. *Brain Lang* 1996; **53:**143-168.

Dogil G, Ackermann H, Grodd W. The speaking brain: a tutorial introduction to fMRI experiments in the production of speech, prosody and syntax. *Journal of Neurolinguistics.* 2002; 15: 55-90.

Donovan A, Darwin CJ. *"The Perceived Rhythm of Speech", in Proceedings of the Ninth International Congress of Phonetic Sciences II*, Eli Fischer-Jorgensen, Jorgen Rischel, and Nina Thorsen (Eds.) pp. 628-74. Copenhagen: University of Copenhagen; 1979.

Dronkers NF, Pinker S, Damasio A. Language and the Aphasias. In Eric R. Kandel, James H. Schwartz, and Thomas M. Jessell. (Eds.) *Principles of Neural Science.* 4th edition, pp. 1169-1187. New York: McGraw-Hill; 2000.

Edmonson J, Chan J-L, Seibert GB, Ross ED. The effect of right-brain damage on acoustical measures of affective prosody in Taiwanese patients. *Journal of Phonetics.* 1987; 15: 219-233.

Eimas P. Auditory and linguistic processing of cues for place of articulation by infants. *Perception and Psychophysics*. 1974; 16: 513-521.

Eimas P. Auditory and phonetic coding of the cues for speech: Discrimination of the [r-1] distinction by young infants. *Perception and Psychophysics*. 1975; 18: 341-347.

Eimas P, Miller J. Contextual effects in infant speech perception. *Science*. 1980; 209: 1140-1141.

Eimas P, Siqueland E, Jusczyk P, Vigorito J. Speech perception in infants. *Science* 1971; 171: 303-306.

Emmorey K. The neurologic substrates for the prosodic aspects of speech. *Brain and Language* 1987; 30: 305-320.

Escera C. Involuntary attention and distractibilioty as evaluated with event-related brain potentials *Audiol. Neuro-Otol.* 2000; 5: 151-166.

Ettlinger G, Jackson C, Zangwill O. Dysphasia following right temporal lobectomy in a right-handed man. *J Neurol Neurosurg Psychiatry.* 1955; 18: 214-217.

Eulitz C, Diesch E, Pantev C, Hampson S and Elbert T. Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *J. Neurosci* 1995; **15:**2748-2755.

Fernandez-Martin F, Martinez-Lage JM, Madoz P. La afasia cruzada: Estudio clinicocon Comprobacion anatomica de un caso. *J Neurol Sci.* 1968; 7: 565-570.

Firth JR, *"Sounds and prosodies".* TPS. 1948; 127-52.

Firth JR, "Sounds and prosodies". In F.R. Palmer (Ed.), *Prosodic Analysis*, pp. 1-26. London: Oxford University Press; 1970.

Friederice AD. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences.* 2002; 6(2): 78-84.

Frith CD. A PET Study of word finding. *Neuropsychologia.* 1991; 29:1137-48.

Fry DB. Prosodic phenomena. In Mallberg B. (Ed.), *Manual of Phonetics*, pp. 365-410. North-Holland Press Amsterdam; 1968.

Gandour J. Aphasia in tone language. In P. Coppens, A. Basso and Y. Lebrun (Eds.), *Aphasia in a typical populations*, pp. 117-141. Hellsdale, NJ: Lawrence Erlbaum; 1998.

Gandour J. Frontiers of Brain Mapping of Speech Prosody. *Brain and Language.* 2000; 71: 75-77.

Gandour J, Dardarananda R. Voice Onset Tome in Aphasia: Thai II Production. *Brain and Language.* 1984; 23: 177-205.

Gandour J, Carney A, Nimitbunnasarn C, Amatyakul P. Tonal confusions in Thai patients with sensorineural hearing loss. *Journal of Speech and Hearing Research.* 1984; 27: 89-97.

Gandour J, Holasuit Petty, S, Dardarananda R. Perception and Production of Tone in Aphasia. *Brain and Language*. 1988; 35: 201-240.

Gandour J, Holasuit Petty S, Dardarananda R. Dysprosody in Broca's Aphasia: A Case Study. *Brain and Language* 1989; 37: 232-257.

Gandour J, Ponglorpisit S, Khunadorn F. Timing characteristics of speech after brain damage: Vowel length in Thai. *Brain and Language*. 1992; 42: 337-345.

Gandour J, Ponglorpisit S, Khundorn F. Stop voicing in Thai after unilateral brain damage. *Aphasiology.* 1992; 6: 535-547.

Gandour J, Ponglorpisit S, Khundorn F. Lexical tones in Thai after unilateral brain damage. *Brain and Language*.1992; 43: 275-307.

Gandour J, Dechongkit S, Ponglorpisit S. Intraword timing relations in Thai after unilateral brain damage. *Brain and Language*. 1993; 45: 160-179.

Gandour J, Dechongkit S, Ponglorpisit S. Speech timing at the sentence level in Thai after unilateral brain damage. *Brain and Language*. 1994; 46: 419-438

Gandour J, Larsen J, Dechongkit S. Speech prosody in affective contexts in Thai Patients with right hemisphere lesions. *Brain and Language*. 1995; 51: 422-443.

Gandour J, Potisuk S, Ponglorpisit S. Tonal coarticulation in Thai after unilateral brain damage. *Brain and Language*. 1996; 52: 505-535.

Gandour J, Wong D, Hsieh L. A Crosslinguistec PET Study of Tone Perecption. *Journal of Cognitive Neuroscience*. 2000; 12 (1): 207-222.

Gao S, Benson DF. Aphasia after stroke in native Chinese speakers. *Aphasiology*. 1990; 4: 31-43.

Goodglass H. *Understanding aphasia*. San Diego: Academic Press; 1993.

Gloning I. Comparison of verbal behaviour in right-handed and non-right-handed patients with anatomically vertified lesions of one hemispheres. *Cortex* 1969; 5: 43-52.

Giard M-H, Perrin F, Pernier J, Bouchet, P. Brain generators implicated in processing of auditory stimulus deviance: a topographic event-related potential study. *Psychophysiology* 1990; 27: 627-640.

Giard MH., Lavikainen J, Reinikainen K, Perrin F, Bertrand O, Thevenet M, Pernier J, Näätänen R, Separate representation of stimulus frequency intensity and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *J. Cognit. Neurosci.* 1995; 7: 133-143.

Goodglass H, Kaplan E. *The assessment of aphasia and related disorders*. Philadelphia: Lea and Febiger. 2nd edition; 1983.

Goodglass H, Quadfasel FA. Language laterality in left-handed aphasics. *Brain* 1954; 77: 523-48.

Goldsmith J. *Autosegmental phonology*. New York: Garland Press; 1979.

Goldsmith J. *Autosegmental and metrical phonology*. Cambridge, MA: Basil Blackwell; 1990.

Halle M, Vergnaud JR. Three-dimensional phonology. *Journal of Linguistic Research.* 1980; 1: 83-105.

Hecaen H, Mzars G, Ramier, A. M. Aphasie croisee chez un sujet droitier bilingue (vietnamien-francais). *Rev Neurol* 1971; 124: 319-323.

Heilman KM., Scholes R, Watson, R.T. Auditory affective agnosia disturbed comprehension of affective speech. *J. Neurol. Neurosurg. Psychiatry.* 1975; 38: 69-72.

Henderson EJA. "Prosodies in Siamese" Asia Major (New Series) I, 1949; 189-215.

Henderson EJA. Prosodies in Siamese. In F.R. Palmer (Ed.) Prosodic Analysis, pp. 27-53 London: Oxford University Press; 1970.

Hermann MJ, Römmler J, Ehlis A-C, Heidrich A. and Fallgatter AJ. Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). *Cogn Brain Res* 2004; 20:294-299.

Hu Y-H, Qiu Y-G, Zhong G-Q. Crossed aphasia in Chinese: A clinical survey. *Brain and Language.* 1990; 39: 347-356.

Hu Y-H, Qiu Y-G, Zhong G-Q. A reply to Gitterman and Sier "Nonbiological determinants of the organization of language in the brain: A comment on Hu, Qiu, and Zhong." *Brain and Language*. 1992; 43: 764-767.

Hughes CP, Chan JL, Ming S-S. Aprosodia in Chinese patients with right cerebral hemisphere lesion. *Archives of Neurology*. 1983; 40: 732-736.

Hung C-C, Tu Y-K, Chen S-H, Chen R–C. A study on handedness and cerebral speech dominance in right-handed Chinese. *Journal of Neurolinguistica*. 1985; 1: 143-163.

Holmes JE, Sadoff RL. Aphasia due to a right hemisphere tumor in a right-handed man. *Neurology*. 1966; 16: 392-397.

Hsieh L, Gandour J, Wong D, Hutchins GD. Functional Heterogeneity of inferior Gyrus is Shaped by Linguistic Experience. *Brain and Language* 2001; 10: 1-15.

Inouchi M, Kubota M, Ferrari P, Toberts, T. Neuromagnetic auditory cortex responses to duration and pitch changes in tones: cross-linguistic comparisons of human subjects in directions of acoustic changes. *Neuroscience Letter* 2002; 331:138-142.

Inouchi M, Kubota M, Ferrari P, Toberts, T. Magnetic mismatch Fields elicited by vowel duration and pitch changes in Japanese words in Humans: comparison between native- and non-speakers of Japanese. *Neuroscience Letter* 2003; 353:165-168.

Inouchi M, Kubota M, Ohta K, Shirahama Y, Takashima A, Horiguchi T, Matsuchima E. Human auditory evoked mismatch field amplitudes vary as a function of vowel duration in healthy first-language speakers. *Neuroscience Letter* 2004; 306:342-346.

Ivry R, Roberson L. The two sides of perception Cambridge, MA: MIT Press; 1998.

Jackson HJ. On affections of speech from disease of the brain. *Brain.* 1915; 2: 203-222.

Jäncke L, Wustenberg T, Scheich H and Heinze H-J. Phonetic perception and the temporal cortex. *Neuroimage* 2002; **15:**733-746.

Jones SJ, Byrne C. The AEP T-complex to synthesised musical tones: left-right asymmetry in relation to handedness and hemisphere dominance. Electroenceph. *Clin. Neurophysiol.* 1998; 108: 355-360.

Jones SJ, Longe O, Vaz Pato, M. Auditory evoked potentials to abrupt pitch and timbre change of complex tones: electrophysiological evidence of "streaming"? *Electroenceph. Clin. Neurophysiol.* 1998; 108: 201-220.

Jusczyk P. The discovery of spoken language. Cambridge MA: MIT Press; 1997.

Kasai K, Yamada H, Kamio S, Nakagome K, Iwanami A, Fukuda M, Itoh K, Koshisa I, Yumoto M, Iramina K, Kato N and Ueno S. Brain lateralization for mismatch response to across-and within-category change of vowels. *Neuroreport* 2001; **12:**2467-2471.

Kent RD, Read C. *The Acoustic Analysis of Speech.* Singular Publishing Group: Sandiego; 1992.

Kent RD, Rosenbeck JC. Prosodic disturbances and neurologic lesion. *Brain and Language.* 1982; 15: 259-291.

Kadota S. The Role of Prosody in Silent Reading. *Language Science*. 1987; 9 (2): 185-206.

Kane NM, Curry SH, Butler SR, Gummins BH, Electrophysiological indicator of awaking from coma. *The Lancet.* 1993; 341: 688.

Kimura D. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*. 1961; 15: 166-171.

Klouda G, Robin D, Grff-Radford N. The role of callosal connections in speech prosody. *Brain and Language*. 1988; 35: 154-171.

Kluender KR, Diehl R, Killeen PR. Japanese quail can learn phonetic categories. *Science.* 1987; 237: 1195-1197.

Knecht K. Language lateralization in healthy right-handers. *Brain*. 2000; 123: 74-81.

Knosche TR, Maess B and Friederici AD. Processing of syntactic information monitored by brain surface current density mapping based on MEG. *Brain Topogr* 1999; **12(2):**75-87.

Kohonen T, Hari R. Where the abstract feature maps of the brain might come from. *Trends in Neuroscience*. 1999; 22: 135-139.

Kraus N, Micco AG, Koch DB. The mismatch negativity cortical evoked potential elicited by speech in cochlear-implant users. *Hear Res*. 1993; 65: 118-124.

Kraus N, McGee T, Carell T, Zecker S, Nicol T, Koch D. Auditory neuro physiologic response and discrimination deficits in children with learning problems. *Science*. 1996; 273: 971-973.

Kuhl PK. Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics.* 1991; 50: 93-107.

Kuhl PK. *Perception cognition and the ontogenetic and phylogenetic* (pp. 73-106). Cambridge, MA: MIT Press; 1991.

Kuhl PK. Learning and representation in speech and language. *Current Opinion in Neurobiology.* 1994; 4: 812-822.

Kuhl PK. Language, mind, and brain: Experience alters perception. In M. Gazzaniga (Ed.), *The cognitive neurosciences* 2nd ed., (pp. 99-115). Cambridge, MA: MIT Press; 2000.

Kuhl PK, Miller JD. Speech perception by the chinchilla: voiced- voiceless distinction in alveolar plosive consonants. *Science.* 1975; 190: 69-72.

Kuhl PK, Padden DM. Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception and Psychophysics*. 1982; 32: 542-550.

Kuhl PK, Padden DM. Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America*. 1983; 73: 1003-1010.

Kuhl PK, Williams KA, Lacerda F, Stevens KA, Lindblom B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*. 1992; 255: 606-608.

Kujala T, Mylly Viita K, Tervaniemi M, Alho K, Kallio J, Näätänen R. Basic auditory dysfunction in dyslexia pinpointed by train-activity measurements. *Psychophysiolo. Special Rep.* 2000; 37: 1-5.

Kutas M, Van Petten C, Besson M. Event-related potential asymmetries during the reading of sentences. *Electroenceph. Clin. Neurophysiol*. 1988; 69: 218-233.

Lancaster JL, Rainey LH, Summerlin JL, Freitas CS, Fox PT, Evans AC, Toga AW, Levanen S, Ahonen A, Hari R, McEvoy L, Sams M. Deviant auditory stimuli activate human left and right auditory cortex differently. *Cereb. Cortex.* 1996; 6: 288-296.

Mazziotta JC. Automated labeling of the human brain – a preliminary report on the development and evaluation of a forward transform method. *Hum Brain Mapp* 1997; 5: 238-242.

Lancaster JL, Summerln JL, Rainey L, Freitas CS, Fox PT. The Talairach Daemon, a database server for Talairach Atlas Labels. *Neuroimage.* 1997; 5 (4): S633.

Lancaster JL, Rainey LH, Summerlin JL, Freitas CS, Fox PT, Evans AE, Toga AW, Mazziotta JC. “Automated Labeling of the Human Brain: A Preliminary Report on the Development and Evaluation of a Forward-Transform Method”, *Human Brain Mapping.* 1997; 5: 238-242.

Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. “Automated Talairach Atlas labels for functional brain mapping” *Human Brain Mapping*. 2000; 10: 120-131.

Lebrun Y, Lessinnes A, De Vresse L. Dysprosody and the non-dominant hemisphere. Language *Science.* 1985; 7: 41-52.

Ledefoeged P. *A Course in Phonetics*. Harcourt Brace Jovanich: New York; 1975.

Ledefoeged P. *A Course in phonetics*. New York: Harcourt Brace Jovanovich. 2nd ed; 1982.

Lehiste I. *Suprasegmentals.* Cambridge, NA: MIT Press; 1970.

Lehiste I. Suprasegmental features of speech. In N. Lass (Ed.) *Contemporary issues in experimental phonetics.* New York: Academic Press; 1976.

Lehiste I. *"Moderator's Introduction (Summary),"* in Proceedings of the Ninth International Congress of Phonetic Sciences II, pp. 241-4, Eli Fischer-Jorgensen, Jorgen Ristchel, and Nina Thorsen (Eds.), Copenhagen: University of Copenhagen; 1979.

Lehmann D. Principles of spatiaal analysiss. In: A.S. Gevins and A. Remond (Eds.) Handbook of Electroencephalography and Clinical Neurophysiology, Rev. Ser., Vol. L: *Methods of Analysis of Brain Electrical and Magnetic Signals*, Elsevier, Amsterdam, pp. 309-354; 1987.

Lehmann D, Skrandies W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysilo.* 1980; 48: 609-62.

Levelt W. *Speaking: From intention to articulation.* Cambridge, MA: MIT Press; 1989.

Levelt W. Producing spoken language: A blueprint of the listener. In C. Brown and P. Hagoort (Eds.) *The neurocognition of language*. pp. 83-122. New York: Oxford University Press; 1999.

Levine RA, MeGaffigan PM. Right left asymmetries in the human brain stem: auditory evoked potentials. Electroenceph. Clin. Neurophysiol. 1983; 55: 532-537.

Levy J. Psychobiological implications of bilateral asymmetry. In: Dimond SJ. Beaumont JG. Editors. *Hemisphere function in the human brain*. New York: Wiley; 1974.

Lewis A. *Infant Speech: A Study of the Beginnings of Language*. Hartcourt Brace: New York; 1936.

Lieberman P, Blumstein S. *Speech physiology, speech perception, and acoustic phonetics.* Cambridge: Cambridge University Press; 1988.

Lind JC, Flor-Henry P. and Koles ZJ. Discriminant analysis and equivalent source localization of the EEG related to cognitive functions. *Brain Topogr* 1999; **11**:265-278.

Mäkelä AM, Alku P and Tiitinen H. The auditory N1m reveals the left-hemispheric representation of vowel identity in humans. *Neurosci Lett* 2003; **353:**111-114.

Marinesco G, Grigoresco D, Axente S. Considerations sur l'aphasie croisee. *Encephale.* 1938; 33: 27-46.

Mateer CA. Neural Correlates of Language Function. In DP. Kueh, ML. Leme, and J.M. Baumgartner. (Eds.), *Neural Bases of Speech, Hearing, and Language*, pp. 259-291 Massachusetts: College-Hill Press; 1989.

Mateer CA, Dodrill CB. Neuropsychological and linguistic correlates of atypical language lateralization: evidence from Sodium Amytal studies. *Hum. Neurobiol*. 1983; 2: 135.

Martin JG. "Rhythmic (Hierarchical) versus Serial Structure in Speech and Other Behavior," *Psychological Review.* 1972; 79: 487-509.

Mattys S. The perception of primary and secondary stress in English. *Perception and Psychophysics.* 2000; 62: 253-265.

Mecaeei I, Spenelli D, Viggiane MP. The affects of visual field size on hemispheric asymmetry of pattern reversal visual evoked potentials. *Int. J. Neurosci.* 1990; 51: 141-151.

Miller J. On the internal structure of phonetic categories: a progress report. *Cognition* 1994; 50 : 271-285.

Milner B, Branch C, Rasmussen T. Observations on cerebral dominance. In A.V.S. de Reuck & M. O'Connor (Eds.), *Ciba Foundation Symposium on Disorders of Language*, pp. 200-214, London: J & A. Churchill; 1964.

Molfese DL, Linnville SE, Wetzel WF. The ontogeny of brain lateralization for speech and nonspeech stimuli. *Brain and Language*. 1975; 2: 356-368.

Molfese DL, Linnville SE, Wetzel WF. Electrophysiological correlates of handedness and speech perception contrasts. *Neuropsychologia.* 1985; 23: 77-86.

Monrad-Krohn GH. Dysprosody or altered melody of language. *Brain.* 1947; 70: 405-423.

Monrad-Krohn GH. The third element of speech: Prosody and its disorders. In Halpern, L. (Ed.), *Problems in Dynamic Neurology*, pp 101-118. Hebrew University Press: Jerusalem; 1963.

Näätänen R, Gaillard AWK, Mantysaalo S. Early selective effect on evoked potential reinterpreted *Acta Psychologia*. 1978; 42: 313-329.

Näätänen R, Picton TW. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure *Psychophysiology* 1987; 24: 375-425.

Näätänen R, *Attention and Brain Function*. Lawrence Erlbaurn, Hillsdale, 1992.

Näätänen R, Winkler I. The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*. 1999; 125: 826-859.

Naeser M, Borod JC. Aphasia in left-handers. *Neurology*. 1986; 36: 471-488.

Nativ A, Pratt H, Allard F, Frank JS. Brain correlates of hand dominance: the association between peripheral and cerebral asymmetries revisited. *Cortex*. 1992; 28: 435-444.

Obleser J, Eulitz C, Lahiri A and Elbert T. Gender differences in functional hemispheric asymmetry during processing of vowels as reflected by the human brain magnetic response. *Neurosci Lett* 2001; **314:**131-134.

Obleser J and Eulitz C. *Proceedings of the 13th International Conference on Biomagnetism*. Berlin, 2002, pp. 359-361.

Obleser J, Elbert T, Lahiri A and Eulitz C. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Cogn Brain Res* 2003; **15:**207-213

Ojemann GA. Cortical organization of language. *Neuroscience.* 1991; 11(8): 2281-87.

Packard JL. Tone production deficits in non-fluent aphasic Chinese speech. *Brain and Language.* 1986; 29: 212-223.

Pascual RD, Michel CM, Lehmann D. Low Resolution Electromagnetic Tomography: A New Method for Localizing Electrical Activity in the Brain. *International Journal of Psychophysiology.* 1994; 18: 49-65.

Pascual-Marqui RD, Gonzalez SL, Valdes P. and Biscay R. Current source density estimation and interpolation based on the spherical harmonic expansion. *Int J Neurosci* 1988; **43:**237-249.

Peretz T. Asymetrie hemispherique dns les amusies. *Rev. Neurol. (Paris).* 1985; 141: 169-183.

Pisoni DB. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Percept Psychophys* 1973; **13:**253-260.

Phillips C. Levels of representation in the electrophysiology of speech perception. *Cognitive Science.* 2001; 25: 711-731.

Poeppel D, Yellin E, Phillips C, Roberts T, Rowley H, Wezler K, Marantz A. Task-induced asymmetry of the auditory evoked M100 neuromagnetic field by speech sounds *Cognitive Brain Research.* 1996; 4: 231-242.

Polka LJ. Linguistics influences in adult perception of non-native vowel contrasts. *J Acoust Soc Am* 1995; **97:**1286-1296.

Polka L and Bohn O. Asymmetries in vowel perception. *Speech Communication* 2003; **41:**221-231.

Price C, Indefrey P, VanTurennout M. The neural architecture underlying the processing of written and spoken word forms. In C. Brown and P. Hagoo (Eds.), *The Neurocognition of Language*, pp. 211-240. New York: Oxford University Press; 1999.

Rasmussen T, Milner B. The role of early left brain injury in determining lateralization of cerebral speech function. *Ann. N.Y.Acad. Sci.* 1977; 299: 355.

Ritter W. Evidence that the mismatch negativity system works on the basis of objects. *NeuroReport.* 2000; 11: 61-63.

Robin D, Tranel D, Damasio H. Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language.* 1990; 39: 539-555.

Robin D, Klouda G, Hug L. Neurogenic disorders of prosody. In M. Cannito and D. Vagal (Eds.), *Treating disordered speech motor control: For clinicians by clinicians.* Austin, TX: Pro – Ed. pp.241-271; 1991.

Ross ED. The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology.* 1981; 38: 561-569.

Ross ED. Affective Prosody and the Aprosodias. In Feinberg T.E. and Farah, M. (Eds.), *Behavioral Neurology and Neuropsychology*, pp. 1-17; 1999.

Ross ED, Mesulam M. Dominant language functionals of the right hemisphere? Prosody and emotional gesturing. *Archives of Neurology.* 1979; 36: 144-148.

Ross ED, Edmondson JA, Seibert GB. The effect of affect on various acoustic measures of prosody in tone and non-tone languages: A comparison based on computer analysis of voice. *Journal of Phonetics*. 1986; 14: 283-302.

Ross ED, Edmonson J, Seibert BB, Chan J-L. Affective exploitation of tone in Taiwanese: An acoustical study of "tone latitude." *Journal of Phonetics*. 1992; 20: 441-456.

Rugg MD. The effects of handedness on event-related potentials in a rhyme-matching task. *Neuropsychologia*. 1985; 23: 765-775.

Ryalls J, Behrens S. An overview of changes in fundamental frequency associated with cortical insult. *Aphasiology*. 1988; 2: 107-115.

Ryalls J, Reinvang I. Functional lateralization of linguistic tones: acoustic evidence from Norwegion. *Lange and Speech*. 1986; 29: 389-398.

Samuel AG. Phonetic prototypes. Perception and Psychophysics 1982; 31: 307-314.

Samson S, Ztorre RJ. Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia.* 1994; 32: 231-240.

Savela J, Kujala T, Tuomainen J, Ek M, Aaltonen O and Näätänen R. The mismatch negativity and reaction time as indices of the perceptual distance between the corresponding vowels of two related languages. *Cogn Brain Res* 2003; **16:**250-256.

Schlanger BB. The perception of emotionally toned sentences by right hemisphere-damaged and aphasic subjects. *Brain and Language*. 1976; 3: 396-403.

Seeck M, Lazeyras F, Michel, CM, Blnke O, Gericke CA, Ives J, Delavelle J, Golay X, Haenggeli CA, de Tribolet N, Landis T. Non-invasive epileptic focus localization using EEG-triggered functional MRI and electromagnetic tomography. *Electroenceph Clin Neurophysiol*. 1998; 106: 508-512.

Seiberman E, Weingartner H. Hemispheric lateralization of functional related to emotion. Brain and Language. 1986; 5: 322-353.

Seron X. Pantonmine interpretation and aphasia. *Neuropsychologia.* 1979; 17: 661-668.

Sinai A. and Pratt H. High-resolution time course of hemispheric dominance revealed by low-resolution electromagnetic tomography. *Clinical Neurophysiology* 2003; **114:**1181-1188.

Sittiprapaporn W. A *Thai dialect geography of Udonthani province: A tonal study.* Faculty of Graduate Study, Mahidol University; 1997.

Sittiprapaporn W. The tonal system of Cha-am dialect (Huaisaytai): A computer – aided study of its acoustic. *Journal of Language of Culture*. 1998; 17: 33-47.

Shapiro BE, Danly M. The role of the right hemisphere in the control of speech prosody in prepositional and affective contexts. *Brain and Language*. 1985; 25: 19-36.

Snow D. The emotional basis of linguistic and nonlinguistic intonation: implicationas for hemispheric specialization. *Developmental Neuropsychology*. 2000; 17(1): 1-28.

Steinhauer K. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat. Neurosce*. 1999; 2: 191-196.

Stewart MG, Jerger J, Lew HL. Effect of handedness on the middle latency auditory evoked potential. *Am. J. Otol*. 1993; 14: 595-600.

Streeter L. Language perception of 2-month old infants shows effects of birth innate mechanisms and experience. *Nature.* 1976; 259: 39-41.

Studdert-Kennedy M. "Speech Perception". Language and Speech. 1980; 23: 45-65.

Sussman E. Attention affects the organization of auditory input associated with the mismatch negativity system. Brain. Res. 1998; 789: 130-138.

Sussman E. An investigations of the auditory streaming effect using event-related brain potentials. *Psychophysiology.* 1999; 36: 22-34.

Talairach J, Tournoux P. *Co-planar stereotaxic atlas of the human brain*, Stuttgart. Thieme; 1988.

Tan U. Cerebral somatosensory potentials evoked by posterior tibial nerve stimulation: lateralization and relation to handedness in left-handed normal subjects. *Int. J. Neurosci*. 1989; 49: 303-317.

Tan U, Komsuoglu S, Akgun A. Inverse relationship between the size of pattern reversal visual evoked potentials from the left brain and the degree of left-hand preference in left-handed normal subjects: importance of the left brain. Int. J. Neurosci. 1993; 72: 79-87.

Tervaniemi M, Kujala A, Alho K, Virtanen J, Ilmoniemi RJ and Näätänen R. Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A magnetoencephalographic (MEG) study. *Neuroimage* 1999; **9:**330-336.

Towle VL, Bolanos J, Suarez D, Tan K, Grzesczuk R, Levin DN, Cakmur R, Frank SA, Spire J-P. The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy. *Electroenceph clin Neurophysiol* 1993: 86: 861-866.

Tsushima T, Takizawa O, Sasaki M, Shieake S, Nishi K, Kohno M., Menyuk P, Best C. *Discrimination of English /r-l/ and /w-y/ by Japanese infants 6-12 months: Language-specific developmental changes in speech perception abilities.* Paper presented at the International Conference on Spoken Language. Yokonama, Japan; 1994.

Van Lancker D. Cerebral lateralization of pitch cues in the linguistic signal. *Internat J. Human Communicalt.* 1980; 13: 201-277.

Van Lancker D, Fromkin V. Hemispheric specialization for pitch and tone: Evidence from Thai. *Journal of Phonetica.* 1973; 1: 101-109.

Van Lancker D, Sidtis. The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *Journal of Speech and Hearing Research.* 1992; 35: 963-970.

Van Lancker D. Disambiguation of dystrophic sentences: Acoustic and phonetic cues. *J. Speech Hearing Res.* 1981; 24: 330-335.

Van Strien JW, Liehi R, Bouma A, Bakker DJ. Event-related potentials during word – reading and figure – matching in left-handed and right-handed males and females. *Brain Lang.* 1989; 37: 525-547.

Vihla M and Salmelin R. Hemispheric balance in processing attended and non-attended vowels and complex tones. *Cogn Brain Res* 2003; **16:**167-173.

Volaitis LE, Miller JL. Phonetic prototypes: influence of place of articulation and speaking rate on the internal structure of voicing categories. *Journal of the Acoustical Society of America.* 1992; 77: 1907-1912.

Wechsler AF. Crossed aphasia in an illiterate dextral. *Brain Lang.* 1976; 3: 164-172.

Weintraub S, Mesulam M, Kramer L. Disturbances in prosody: A right hemisphere contribution to language. *Archives of Neurology*. 1981; 38: 724-744.

Werker JF. Cross-language speech perception: developmental change does not involve loss. In H. C. Nusbaum J. Goodman (Eds.) *The development of speech perception: The transition from speech sounds to spoken words* (pp. 93-120). Cambridge, MA: MIT Press; 1994.

Werker JF. Exploring development changes in cross-language spoeech perception. In L. Gleitman M. Liberman (Eds.) *An invitation to cognitive science*, part I: Language (pp. 87-106). Cambridge, MA: MIT, Press; 1995.

Werker JF, Lalone CE. Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*. 1988; 24: 627-683.

Werker JF, Tees RC. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavioural Development* 1984; 7: 49-63.

Werker JF, Tees RC. Influences on infant speech processing: toward a new synthesis. *Annual Review of Psychology*. 1999; 50: 509-535.

Wechsler D. Wechsler *Adult intelligence Scale Manual*. New York: Psychological Co; 1995.

Wildgruber D, Pihan H, Erb M. Differential activation patterns during discrimination of affective prosody: influence of acoustic, emotional valence, accuracy and sex. *NeuroImage.* 1999; 19: s1018.

Winkler I. Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Res*. 1996; 7742: 239-252.

Winkler I, Czigler I. Mismatch negativity: deviance detection or the maintenance of the 'standard' *NeuroReport.* 1998; 6: 3809-3813.

Winkler I, Lehtokoski A, Alko P, Vainio M, Czigler I, Csepe V, Aaltonen O, Raimo I, Alho K, Lang H, Livonen A, Näätänen R. Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*. 1999; 7: 357-369.

Woerner W. *Tone processing and the brain*. In proceedings of the Third International Symposium on Language and Linguistics: Pan-Asiatic linguistics (Vol. 1, pp. 588-603). Bangkok: Chulalongkorn University Press; 1992.

Yabe H. Temporal window of integration of auditory information in the human brain. *Psychophysiology*. 1998; 34: 615-620.

Zatorre RJ, Hepern AR. Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia*. 1993; 31: 221-232.

Zatorre RJ, Samson S. Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*. 1991; 114: 2403-2417.

Zatorre RJ, Evans A, Meyer E. Cortical mechanism underlying melodic perception and memory for pitch. *Journal of Neuroscience*. 1994; 14: 1908-1919.

Zatorre RJ. Auditory attention to space and frequency activates similar cerebral systems. *Neurolmage*. 1999; 10: 544-554.