Exploring the effects of accent on cognitive processes: behavioral and electrophysiological insights

Doctoral dissertation by:

Trisha Breanne Thomas

Supervised by:

Prof. Clara Martin and **Dr. Sendy Caffarra**

emun ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

2023

Basque Center on Cognition, Brain and Language (BCBL)
Paseo Mikeletegi, 69
Donostia-San Sebastián, Spain
2023

The work presented in the current dissertation was supported by the Basque Government [BERC 2018-2021 program]; the Spanish State Research Agency [BCBL Severo Ochoa excellence accreditation SEV-2015-0490]; the H2020 European Research Council [ERC Consolidator Grant ERC-2018-COG-819093 to CDM; Marie Sklodowska-Curie grant 837228 to SC]; the Spanish Ministry of Economy and Competitiveness [PID2020-113926GB-I00 to CDM]; the Basque Government [PIBA18-29 to CDM]; the Italian Ministry of University and Research (Programma giovani ricercatori Rita Levi Montalcini) to SC; and the Programa Predoctoral de Formación de Personal Investigador No Doctor del Departamento de Educación del Gobierno Vasco (PRE_2021_2_0006 to TT). Support was also provided by the Basque Government short stay grant to conduct a three-month research stay at the Digital Learning Lab at the University of California Irvine, under the supervision of Dr. Mark Warschauer on the Converse to Learn Project. The study carried out during the research stay was not included in the main experimental chapters of the dissertation but is summarized in chapter 7.1.

Exploring the effects of accent on cognitive processes: behavioral and electrophysiological insights

Doctoral dissertation by:

Trisha Breanne Thomas

Supervised by:

Prof. Clara Martin and **Dr. Sendy Caffarra**

A dissertation submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy in Cognitive Neuroscience

Basque Center on Cognition, Brain and Language (BCBL)

Universidad del País Vasco / Euskal Herriko Unibersitatea (UPV/EHU)

Donostia-San Sebastián, Spain, 2023

emman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

BASQUE CENTER ON COGNITION, BRAIN AND LANGUAGE

Acknowledgements

Thank you to everyone who supported, guided, and challenged me throughout these past four years. I could not have done this without my community and all of the work would be meaningless without you all. First and foremost, I owe infinite gratitude to my supervisors Sendy Caffarra and Clara Martin, who were always available when I needed guidance and not only are two excellent female scientists and role models but also wonderful human beings. A special thank you to Sendy who frequently woke up early while in California in order to have meetings with me because of our 9-hour time difference.

Thank you to all the admin, IT, and lab team at BCBL who helped make these four years run smoothly. Thank you also to the Basque Government for their generous funding of my project. Thank you to all my friends and colleagues from the PhD at BCBL, the pre-doc group and my 'remaining snoggers' masters family-Jessi, Catherine, Kirk, Meng Xing, Magda, Mina, and Pratik (from afar), with a special thank you Dr. Meng Xing Liu who graciously dealt with many hours of me asking him for programming help and Jessi and Catherine who were my emotional pillars. I am so excited to see where all of your paths lead you both as successful scientists and in your personal adventures. Thank you also to the Digital Learning Lab at the University of California Irvine for hosting me during my research stay and providing me with an invaluable experience on an exciting project.

To all of my friends, family and loved ones from outside the BCBL, both here and resting, thank you so much for always supporting me. It would be impossible to list the names of all those to whom I owe my gratitude so I will reserve this section only for the Woolridge family. I want to include a special thank you to Joseph Wooldridge who, while no longer with us, first encouraged me to consider going abroad to pursue a master's degree and to Don Wooldridge (and my dog nephew Uri) who pushed me not to cancel my plans to go after we lost Joe and then supported me the whole way through our hours of video call boardgame sessions. I was so lost when Joe passed and I know would not be here today without you both.

Finally, thank you so much to my parents, my siblings, my nieces and nephew, and my grandparents who have supported me all the way throughout my education and always encouraged me to follow my academic ambitions, even when they led me far away from home. Thanks, Mom and Dad, for always doing anything in your power to facilitate my ability to follow my aspirations. I want to thank all the strong, rebellious women in my family who never

let anyone hold them back in the pursuit of their dreams, my late Grandmothers Eileen and Doreen, my mom, my sisters Jen and Nina, my niece Mirah who is firmer willed at 7 than I was at 18. I couldn’t finish this section without a special mention to my Grandmother, Doreen, who called to check in on my Ph.D. progress multiple times a week every week for years and has been looking forward to this moment as much as me, if not more. My grandma grew up on a farm, had to push to continue schooling past 14 and valued education almost as much as she valued service to others. Grandma really wanted to be around to call me Dr. Thomas and tried her best to hang on near the end to do so. Unfortunately, she passed away just before I finished the dissertation. I wish I could have had that phone call, finally answering her cheerful ‘how’s the paper coming?’ with the good news of it’s completion but life rarely gives us the closure we are looking for. I want to thank her especially because it was her Welsh lessons as a child that first sparked my interest in multilingualism, language acquisition and minority languages. I do not think I would be here today if not for my early years learning Welsh and going to Welsh cultural events with my grandparents. Diolch yn fawr iawn nain, wedi'i wneud yn olaf.

Resumen en Castellano

Capítulo 1

En primer lugar, en la introducción general explico cómo se había estudiado el procesamiento del habla anteriormente y cómo las innovaciones tecnológicas y científicas han hecho avanzar este campo. A continuación, expongo cómo el procesamiento del habla puede verse afectado por pistas dependientes del contexto y, por último, analizo el papel específico del acento del hablante.

El habla es una forma de comunicación que comprende la producción y percepción de sonidos según las reglas de una lengua. El habla depende de varios articuladores del aparato fonador, pero estos movimientos son coordinados y, en el caso de la percepción, descifrados por el cerebro. Así pues, el procesamiento del habla puede describirse a grandes rasgos como la abstracción satisfactoria del significado a partir de una señal acústica. La extracción de información de la señal requiere de una serie de transformaciones de la señal acústica a medida que viaja desde el sistema perceptivo auditivo hasta las áreas de procesamiento del cerebro.

El habla difiere de la entrada auditiva primaria debido a esta compleja interacción entre el sistema auditivo y el sistema del lenguaje. El habla puede analizarse en varios niveles. El nivel fonológico del análisis del habla se centra en los sonidos del lenguaje y en cómo se utilizan para formar palabras. El nivel léxico-semántico del análisis del habla se ocupa del significado de las palabras y de cómo se combinan los sonidos para formar unidades de significado más amplias. Por último, el análisis oracional y discursivo se centra en la estructura de las frases y en el modo en que estas se combinan para formar conversaciones. La presente investigación se centrará en estos tres niveles.

Los científicos llevan siglos estudiando el procesamiento del habla; el estudio metodológico de la producción y percepción del habla comenzó sobre todo con la observación fonológica (véase Casserly & Pisoni, 2010) y el análisis patológico a través de casos de pacientes. Entre 1920 y 1930, comenzaron las primeras grabaciones de la actividad eléctrica para describir las actividades del habla en el cerebro. El estudio del Potencial Evocado Auditivo (PEA) se remonta a finales de la década de 1940 y principios de la de 1950, cuando los investigadores empezaron a investigar la actividad eléctrica del cerebro en respuesta a estímulos auditivos, ya que proporcionan una forma de evaluar la integridad del sistema auditivo y la respuesta del cerebro al sonido. El PEA sigue siendo una herramienta valiosa para el estudio del procesamiento auditivo y el diagnóstico de los trastornos auditivos. Además del PEA, los científicos también empezaron a utilizar técnicas como la EEG y, posteriormente, la MEG para relacionar las frecuencias de las ondas cerebrales y los PRE con otros aspectos específicos del procesamiento del habla, como la extracción acústica, la normalización entre la entrada acústica y las representaciones fonológicas y la integración léxico-semántica.

Junto con los avances electrofisiológicos, a mediados del siglo XX surgió el movimiento conductista de la psicología. El conductismo daba prioridad al estudio de los comportamientos observables, frente a procesos mentales internos como la percepción y la cognición. La influencia de este movimiento dio lugar a uno de los primeros modelos de percepción del habla, la Teoría Motora. El conductismo también influyó en el campo de las primeras investigaciones sobre el procesamiento del habla, que se centraban principalmente en las propiedades acústicas del habla y en las respuestas del oyente a los sonidos. Aunque ahora los científicos son conscientes de la importancia de los mecanismos cognitivos en el procesamiento del habla, los experimentos centrados en medidas conductuales, como los tiempos de reacción, siguen contribuyendo de forma decisiva al creciente conocimiento en este campo.

El uso de modelos computacionales en la investigación del habla comenzó en torno a la década de 1960 y aumentó hacia las décadas de 1980 y 1990, a medida que se producían avances en la tecnología y los algoritmos. El Modelo Oculto de Markov (HMM), desarrollado en las décadas de 1970 y 1980, fue uno de los primeros y más importantes modelos informáticos en la investigación del procesamiento del habla (véase (Kayte et al., 2015). Los HMM modelan secuencias de vectores espectrales variables en el tiempo y aún se utilizan como base de los sistemas de reconocimiento del habla continua.

Por último, técnicas de neuroimagen como la resonancia magnética funcional (fMRI) y la tomografía por emisión de positrones (PET) se utilizaron de forma generalizada para estudiar el procesamiento del habla en las décadas de 1990 y 2000. Estas técnicas aumentaron la resolución espacial del conocimiento del cerebro, lo que permitió a los científicos observar qué regiones cerebrales estaban activas durante distintos aspectos del procesamiento del habla y actualizar modelos neurobiológicos anteriores creados a partir de casos de pacientes.

Gracias a la combinación de las técnicas descritas anteriormente, ahora tenemos una visión bastante precisa del viaje del sonido hasta el cerebro. El procesamiento del habla comienza cuando las ondas sonoras son interceptadas por el oído. Las ondas sonoras entran en el oído y hacen vibrar el tímpano; estas vibraciones continúan hasta el oído interno, donde la señal física se convierte en señal eléctrica. A continuación, estas señales eléctricas viajan desde la cóclea hasta el nervio auditivo y el tronco encefálico, donde se procesan y filtran. Una vez procesadas en el tronco encefálico, las señales auditivas se envían al tálamo, que actúa como estación de retransmisión de la información sensorial. A continuación, el tálamo envía las señales auditivas al córtex auditivo primario, situado en el lóbulo temporal, a través de un haz de fibras nerviosas denominado radiación acústica. Una vez que las señales auditivas llegan a la corteza auditiva primaria, se procesan e integran con otra información sensorial en el cerebro para dar lugar a nuestra percepción del sonido y se propagan a otras regiones cerebrales implicadas en el procesamiento del habla, como la circunvolución temporal superior, el área de Broca y el área de Wernicke. A medida que se procesa la información del habla, se integra para formar una representación coherente del habla. Las características lingüísticas se combinan para formar una interpretación significativa. A lo largo de los años, se han propuesto varios modelos para explicar cómo procesa el cerebro el habla. Ahora sabemos que múltiples áreas del córtex desempeñan funciones importantes en el procesamiento del habla y del lenguaje. Los primeros modelos de percepción del habla pretendían entender cómo el cerebro procesa y comprende el habla. Dos modelos opuestos en este campo fueron el Enfoque Realista Directo y la Teoría Motora. El Enfoque Realista Directo (Fowler, 1986) plantea que la percepción del habla implica el procesamiento directo de señales acústicas en el sistema auditivo sin necesidad de representaciones intermedias. La Teoría Motora de la percepción del habla (Liberman & Mattingly, 1985) propone que la percepción y la producción del habla están estrechamente entrelazadas, y que el cerebro utiliza representaciones motoras para procesar los sonidos del habla.

En los últimos años se han propuesto varios modelos para explicar los mecanismos subyacentes de la percepción del habla. El modelo de doble flujo y el modelo TRACE representan marcos destacados para comprender la percepción del habla desde una perspectiva neurobiológica. Mientras que el modelo de doble flujo hace hincapié en los distintos flujos de procesamiento y en la integración contextual, el modelo TRACE se centra en la interacción de las representaciones fonémicas, léxicas y semánticas. Estos modelos aportan valiosos conocimientos sobre los mecanismos neuronales y los procesos cognitivos que intervienen en la percepción del habla, contribuyendo a una comprensión más profunda de cómo los seres humanos perciben y comprenden el habla.

Numerosos modelos de procesamiento del habla dependientes del contexto han hecho hincapié en la importancia de las señales indiciarias a la hora de transmitir significado. Se trata de señales contextuales relacionadas con el hablante, como sus gestos, expresiones faciales, tono de voz y otras características que pueden utilizarse para transmitir información. El modelo episódico hace hincapié en el papel de los rastros episódicos almacenados, mientras que el modelo estándar de dos pasos hace hincapié en la integración de los indicios indexados con otros conocimientos lingüísticos en el proceso de mapeo. En conjunto, estos modelos aportan información valiosa sobre la compleja interacción entre la información acústico-fonética y el conocimiento lingüístico de alto nivel, incluidas las claves indiciales, en la percepción del habla.

El acento es una cualidad indicial del habla que tiene cada persona y que viene determinada en gran medida por el entorno geográfico, lingüístico y cultural en el que hemos aprendido una lengua. Un acento, en su sentido más amplio, se refiere a la forma en que una persona habla un idioma. Es importante señalar que, a los efectos de la investigación llevada a cabo en esta tesis, voy a clasificar el acento en dos grandes categorías y a subdividir una de ellas. En concreto, utilizaré los términos nativo, extranjero y dialectal para reflejar las variaciones del acento nativo (nativo y dialectal) y no nativo (extranjero). Los retos que plantea el acento no nativo al oyente nativo no son solo lingüísticos, sino también cognitivos en un sentido más amplio. Los retos lingüísticos analizados en este capítulo son: fonológicos y léxico-semánticos. Los retos cognitivos analizados en este capítulo son: atencionales y de memoria. También se analizan los retos asociados a la integración de los efectos lingüísticos y cognitivos, así como los marcos teóricos de todos los retos.

Capítulo 2

En este capítulo se presentará la metodología de EEG, seguida de los hallazgos conductuales y electrofisiológicos sobre el procesamiento del acento. Dado que el habla es un medio que se desarrolla muy rápidamente, la información a una escala temporal muy sensible es útil para comprender mejor lo que ocurre durante la percepción y el procesamiento, en contraposición a los efectos conductuales del procesamiento únicamente.

La electroencefalografía es una herramienta de neuroimagen no invasiva que mide la actividad neuronal en el cuero cabelludo. Es una herramienta especialmente útil para el estudio de la información temporal detallada de los procesos del lenguaje en el cerebro. El EEG se registra mediante múltiples electrodos que se colocan en el cuero cabelludo del participante y permiten seguir las fluctuaciones de la actividad electrofisiológica cerebral. Gracias a la EEG, los investigadores pueden aislar y localizar temporalmente los correlatos neuronales de una serie de procesos específicos relacionados con el lenguaje a una escala de milisegundos. Los análisis de dominio temporal, como los potenciales relacionados con eventos (PRE), son un conjunto de medidas de EEG muy populares. Los PRE son las medidas promediadas de voltajes extremadamente pequeños que se producen cuando un gran número de neuronas se disparan simultáneamente en respuesta a un único evento interno o externo (por ejemplo, un movimiento ocular o la presentación de una palabra hablada durante un experimento, respectivamente). Los componentes comunes de los ERP incluyen el P2, el PMN y el N400 (que se analizará en el capítulo 3), entre otros (Sur y Sinha, 2009; Luck, 2014). Otro análisis EEG consiste en la descomposición tiempo-frecuencia. Con este método, también se tiene en cuenta el patrón oscilatorio de la actividad cerebral. La descomposición de la señal cerebral mediante técnicas matemáticas basadas en el análisis de Fourier proporciona indicadores de actividad oscilatoria. Este enfoque permite descomponer la señal cerebral en muchos componentes en función de varias gamas de frecuencia, es decir, oscilaciones neuronales, en lugar de reducir la señal cerebral a una única variable. Las bandas de frecuencia suelen dividirse en cinco categorías: delta (1-4 Hz), theta (4-8 Hz), alfa (9-12 Hz), beta (13-30 Hz) y gamma (> 30 Hz). Ambas técnicas son complementarias y permiten estudiar tanto los mecanismos descendentes como los ascendentes. Los retos expuestos en el capítulo anterior han servido de catalizador para multitud de estudios conductuales y de EEG sobre cómo el oyente nativo procesa el habla no nativa. Las pruebas conductuales y EEG se discuten en el contexto de los retos y se contextualizan con los modelos al final de sus respectivas secciones.

Aunque muchos estudios anteriores han demostrado que el acento extranjero afecta al procesamiento del habla y otros tantos han sugerido que puede tener efectos residuales en procesos complejos como las representaciones de la memoria, hay menos consenso sobre en qué medida y cómo cambian esos efectos a lo largo del curso temporal del procesamiento lingüístico. Además, todavía se debate cómo varía el procesamiento dialectal en relación con el acento nativo y el extranjero. Es necesario seguir investigando si los acentos extranjeros se procesan de forma única en comparación con todas las formas del habla nativa, incluyendo tanto el acento nativo como las variaciones dialectales. Queda por determinar si el acento dialectal se percibe de forma diferente al acento nativo, a pesar de ser ambos variaciones del habla nativa. Basándonos en los modelos y pruebas descritos anteriormente, sabemos que aún existen resultados opuestos sobre el efecto de los distintos tipos de acento en el curso temporal del análisis lingüístico y hay al menos tres cuestiones principales de investigación que siguen sin resolverse.

En consecuencia, el objetivo central de esta tesis sigue siendo investigar más a fondo los mecanismos de procesamiento del acento del habla a través de diferentes niveles de análisis lingüístico utilizando pruebas procedentes de experimentos tanto conductuales como electrofisiológicos. Además, divido este tema de investigación en tres preguntas de investigación específicas distribuidas a lo largo de los siguientes capítulos con el fin de abordar los retos mencionados:

Retos lingüísticos:

PI1. ¿Cómo afecta el acento a los mecanismos de procesamiento temprano durante el procesamiento aislado de una sola palabra?

Retos cognitivos:

PI2. ¿Cómo afecta el acento al esfuerzo de escucha y a la memoria del procesamiento de frases?

Retos de integración cognitivo-lingüística:

PI3. ¿Cómo afecta el acento a los mecanismos de procesamiento durante el procesamiento del discurso extendido?

Capítulo 3

El objetivo de este estudio era desentrañar las hipótesis contrapuestas sobre el procesamiento del acento dialectal: la «Hipótesis de los Procesos Diferentes» y la «Hipótesis de la Distancia Perceptiva» y comprender los mecanismos del procesamiento temprano del acento.

Se registraron datos electroencefalográficos de 25 bilingües tempranos de español-euskera del País Vasco que escucharon 40 palabras aisladas en español con diferentes acentos (nativo, dialectal, extranjero). Se emplearon métodos electrofisiológicos por su capacidad para proporcionar información muy valiosa sensible al tiempo. Se extrajeron las amplitudes medias de los potenciales relacionados con eventos: P2 [150-250 ms], PMN [250-400 ms] y N400 [400-600 ms]. Distintas ventanas temporales respaldaron la «Hipótesis de los Procesos Diferentes». Los resultados muestran que los mecanismos de procesamiento tempranos distinguen solo entre el habla nativa y no nativa, con una amplitud P2 reducida para el procesamiento del acento extranjero, lo que apoya la «Hipótesis de los Procesos Diferentes». Además, los mecanismos de procesamiento posteriores muestran una diferencia binaria similar en el procesamiento de los acentos, con una mayor negatividad PMN en el acento extranjero en comparación con los otros, lo que apoya aún más dicha hipótesis.

Nuestros resultados respaldan la idea de que, incluso durante la escucha pasiva de palabras aisladas, los rasgos de indicidad como el acento afectan a nuestro procesamiento en etapas tempranas. Además, nuestros resultados concuerdan con la idea de que los acentos extranjeros se procesan de forma diferente a los acentos dialectales o nativos, lo que requiere un mayor esfuerzo para extraer las características acústicas y normalizarlas en representaciones fonológicas esperadas. A pesar de ello, parece que somos capaces de normalizar con éxito el habla no nativa y, por tanto, en etapas posteriores del procesamiento, el acento ya no afecta a nuestro procesamiento ni a nuestra capacidad de integrar los sonidos del habla en la información léxica.

Capítulo 4

En este estudio examinamos los efectos del acento en el esfuerzo auditivo, en la memoria y en su interacción. Millones de personas interactúan a diario con hablantes no nativos de su lengua. Aunque cada vez es más probable que mantengamos conversaciones con hablantes extranjeros, se ha demostrado que escuchar el habla con acento extranjero nos plantea dificultades únicas (Lev-Ari, 2014, Atkinson et al., 2005). Además, a pesar de saber que el acento puede provocar diferencias tanto en la atención como en las representaciones de la memoria (Levi-Ari, 2017, Atkinson et al., 2005), se sabe poco sobre la posible relación causal entre ellas. Además, se sabe aún menos sobre las consecuencias adicionales del acento extranjero sobre la capacidad para recordar información proporcionada por un hablante no nativo durante un periodo de varios días. El objetivo de este estudio era examinar el efecto de distintos acentos en las demandas cognitivas y la posterior retención de la información.

En base al vacío existente en la literatura relacionada con esta temática, se han propuesto tres modelos diferentes dentro del ámbito de la distribución de recursos atencionales y los efectos posteriores sobre la retención de la memoria: la «Teoría de la Carga Cognitiva», la «Teoría del Esfuerzo Cognitivo» y la «Teoría Libre de Indicidad». Mediante el paradigma de la tarea dual para medir la atención a través del esfuerzo auditivo, planteamos la hipótesis de que las desviaciones del acento nativo en una tarea auditiva aumentarían la demanda de distribución de recursos atencionales (esfuerzo auditivo) y, por tanto, disminuirían el rendimiento en la tarea secundaria. Con el apoyo de la «Teoría de la Carga Cognitiva», se podría predecir que una mayor demanda atencional resultaría en una menor carga de información y un peor rendimiento en la retención de la memoria en el caso de los acentos extranjeros, en comparación con los nativos. Por el contrario, la «Teoría del Esfuerzo Cognitivo» predeciría que una mayor demanda atencional daría como resultado una mejor carga de información y rendimiento en la retención de memoria en el caso de los acentos extranjeros, en comparación con los nativos. Por último, la «Teoría Libre de Indicidad» predeciría que, independientemente de la variación de la demanda atencional, no habría diferencias en el rendimiento de retención de la memoria en el caso delos acentos extranjeros en comparación con los nativos.

Para investigar estas teorías, 36 bilingües tempranos de español-euskera del País Vasco siguieron un paradigma de tarea dual, escuchando frases en español con múltiples acentos (nativo, dialectal, extranjero) en un diseño mediante bloqueo, mientras completaban una tarea secundaria (seguimiento de círculos en la pantalla). Para cada tipo de acento, se recogieron medidas de esfuerzo auditivo y puntuaciones de memoria en una sesión inmediata y en una posterior (4-6 días).

Los resultados muestran que, mientras que un mayor esfuerzo auditivo se asocia con un menor rendimiento de la memoria en el caso del acento nativo, el acento extranjero parece alterar esta compensación entre atención y memoria. De hecho, el acento extranjero parece tener un efecto beneficioso en el rendimiento general de la memoria, independientemente del esfuerzo auditivo. Los resultados sugieren que la relación entre el esfuerzo auditivo y la memoria no parece ser fija, sino que cambia en función del acento.

Capítulo 5

El objetivo de este estudio también era examinar los efectos del procesamiento del acento durante el habla continua y, además, abordar las hipótesis contrapuestas sobre el procesamiento del acento dialectal: la «Hipótesis de los Procesos Diferentes» y la «Hipótesis de la Distancia Perceptiva».

Se registraron datos electroencefalográficos de 29 bilingües tempranos de español-euskera del País Vasco que escucharon 3 diálogos en español de 6 minutos grabados con diferentes acentos (nativo, dialectal, extranjero). En el análisis de los datos se incluyó la estimación de la densidad espectral de potencia (1-80 Hz) y la normalización (restando el bloque de silencio) para cada condición de acento. Nuestro objetivo era descubrir en qué bandas de frecuencia difiere el procesamiento del habla nativa, regional y extranjera y dónde se sitúa el acento regional con respecto al extranjero.

Descubrimos que, mientras que en los rangos de frecuencia más altos el procesamiento del acento extranjero se diferencia de las amplitudes de potencia del procesamiento del acento nativo, en las frecuencias bajas no observamos ninguna modulación de amplitud de potencia relacionada con el acento. Esto sugiere que puede haber una diferencia en el procesamiento de fonemas para el acento nativo y extranjero, mientras que especulamos que los mecanismos descendentes durante el procesamiento del discurso pueden mitigar otros efectos observados con unidades más cortas de discurso.

Capítulo 6

En este capítulo de conclusiones, retomo las preguntas de investigación planteadas en la introducción y las abordo a la luz de los principales resultados de los estudios presentados en este trabajo doctoral. En primer lugar, resumo los resultados de cada estudio y, a continuación, expongo las implicaciones teóricas de estos hallazgos y los sitúo en el contexto más amplio del marco general del procesamiento lingüístico a la luz de los resultados contrastivos sobre los efectos de los distintos tipos de acento en el curso temporal del análisis lingüístico. Los retos lingüísticos fueron el foco de la PI1, en la que he explorado los mecanismos de procesamiento temprano del acento en el nivel fonológico y léxico-semántico del procesamiento de una sola palabra. La PI2 se centró en los retos cognitivos del procesamiento del habla con acento, explorando el esfuerzo auditivo atencional y los efectos en la memoria de la variación del acento a medida que el nivel lingüístico de procesamiento se vuelve ligeramente más complejo. Por último, en la PI3 se exploraron los retos relacionados con la integración del procesamiento cognitivo y lingüístico debido a la interacción de los mecanismos ascendentes y descendentes que permite el curso temporal más largo del discurso fluido.

Los experimentos de esta tesis tenían como objetivo desentrañar una serie de hipótesis sobre el procesamiento de los acentos y el vínculo más amplio entre los procesos de atención y memoria. Además de poner a prueba estas hipótesis, en este capítulo se ha intentado situar el trabajo en el contexto de los marcos teóricos comentados en la introducción. Aunque el modelo de doble flujo y el modelo TRACE ofrecen perspectivas diferentes sobre la percepción del habla, no son mutuamente excluyentes. De hecho, ofrecen perspectivas complementarias sobre distintos aspectos de los procesos perceptivos y cognitivos implicados. El modelo de doble flujo hace hincapié en las vías neuronales y los flujos de procesamiento que intervienen en el análisis de la información acústica y la integración de las claves contextuales, mientras que el modelo TRACE se centra en la interacción de las representaciones fonémicas, léxicas y semánticas. Juntos, estos modelos contribuyen a una comprensión más completa de cómo se percibe y procesa el habla en el cerebro humano. En el contexto del procesamiento del acento del habla, deben tenerse en cuenta ambos modelos. Aunque no son pruebas obtenidas mediante neuroimagen, los resultados combinados de estos experimentos apoyan la plausibilidad de flujos de procesamiento separados para el procesamiento acústico y el léxico-semántico. Efectivamente, parece existir una diferenciación en línea entre los acentos principalmente para el procesamiento acústico sobre el procesamiento de nivel superior. Sin embargo, la compleja interacción entre los niveles de análisis lingüístico y los efectos del acento incluso en procesos cognitivos generales como la atención y la memoria sugieren que también puede existir una red conectiva multidireccional entre los distintos procesos.

En conjunto, este trabajo de doctorado aporta pruebas de que el cerebro humano es muy sensible al acento del habla, incluso en fases muy tempranas de la percepción (capítulo 3), y de que parece haber indicios de que los distintos tipos de acento afectan de manera diferente a los distintos niveles de los procesos cognitivos. Mientras que el acento extranjero parece afectar de forma única a las respuestas evocadas tempranas (capítulo 3) y a los rangos de alta frecuencia (capítulo 5), tanto el acento dialectal como el acento extranjero parecen afectar al rendimiento de la memoria (capítulo 4). Sorprendentemente, tuvieron un efecto beneficioso en el rendimiento de la memoria de reconocimiento, lo que sugiere que es necesario explorar más, especialmente en lo que respecta a cómo afectan los acentos a la retención de información durante largos periodos de tiempo. Aunque el acento afecta a una fase temprana del análisis fonológico, parece resolverse más tarde en la fase semántica. A pesar de ello, esta resolución parece tener efectos persistentes en la memoria. En base a los hallazgos presentados aquí, el estudio del procesamiento del acento en el habla se beneficiaría de considerar la posibilidad de mecanismos de procesamiento flexibles que se adapten a lo largo del curso temporal del procesamiento del acento. Mientras que los modelos de procesamiento del acento dependientes del contexto hacen hincapié en cómo el procesamiento puede adaptarse a un interlocutor con acento no nativo, es posible que esta dependencia del contexto en sí no sea fija y que se trate más bien de una serie de mecanismos cognitivos en constante actualización que se adaptan milisegundo a milisegundo a los acentos, teniendo en cuenta el contexto social, lingüístico y cognitivo de la experiencia y los recursos cognitivos del oyente en tiempo real.

Abstract

With increased international mobility, people are interacting with foreigners at an unprecedented rate. Interlocutors in essential information-disseminating roles-such as our doctors, teachers and co-workers- are increasingly likely to come from a different background than us making it imperative to continue to advance investigations of how accents that differ from our own affect how we process, perceive and remember speech.

Previous research has found that speaker accent can have an impact on a range of offline and online cognitive processes (Baus, Bas, Calabria, & Costa, 2017; McAleer, Todorov, & Belin, 2014; Stevenage, Clarke, & McNeill, 2012; Sporer, 2001). Accordingly, the central aim of this dissertation is to further investigate processing mechanisms of speech accent across different levels of linguistic analysis using evidence from both behavioral and electrophysiological experiments. In particular, I focused on the levels of linguistic information: the word, the sentence and the extended dialogue.

Experiment 1 focused on how accent affects early processing mechanisms during isolated single-word processing, exploring linguistic challenges of speech accent using EEG ERP evidence to differentiate between early, bottom-up processing mechanisms in native, dialectal and foreign accent. Experiment 2 focused on how accent affects the listening effort and memory of sentence processing, exploring the cognitive challenges of attentional and memory effects of dialectal and foreign accent as the linguistic level of processing becomes slightly more complex. Experiment 3 focused on how accent affects processing mechanisms during extended discourse processing, exploring the challenges related to the integration of cognitive and linguistic processing due to the interaction of bottom-up and top-down mechanisms that the longer time course of fluid discourse allows for. The role of various pre-existing models of speech processing is also discussed.

Taken as a whole, this doctoral work provides evidence that the human brain demonstrates a high sensitivity to speech accent, even in the early stages of perception (as discussed in Chapter 3). Furthermore, the findings suggest that various types of accents may have distinct impacts on different cognitive processes. While foreign accent seems to uniquely impact early evoked responses (Chapter 3) and high-frequency ranges (Chapter 5), both dialectal accent and foreign accent appear to affect memory performance (Chapter 4). Surprisingly, they had a beneficial effect on recognition memory performance, suggesting more needs to be explored especially in regards to how accents affect information retention over extended periods of time. Based on the findings presented here, the study of speech accent processing would benefit from considering the possibility of flexible processing mechanisms that adapt along the time course of accent processing.

Table of Contents

Acknowledgements 5
Resumen en Castellano 7
Abstract 18
Prelude **22**
Overview of the work 22
1. Chapter I: General Introduction **23**
1.1. Speech processing: background and rationale 23
1.1.1. Historical perspectives 23
1.1.2. Modern perspectives 26
1.1.3. Role of context in speech processing 31
1.1.3.1. Context-dependent models and indexical cues 32
1.1.3.2. Speaker identity as an indexical cue 33
1.2. Accent perception and processing 33
1.2.1. Accent as a case of speaker identity 33
1.2.1.1. Accent types 34
1.2.1.2. Challenges of foreign-accented speech 35
1.2.1.3. Competing theoretical frameworks of accent perception and processing 38
2. Chapter II: Experimental evidence on accent perception and processing **43**
2.1. Methodology 43
2.1.1. EEG 43
2.1.1.1. Evoked activity time domain 44
2.1.1.2. Oscillatory activity frequency domain 45
2.2. Behavioral studies on accent effects on cognitive processes 46
2.3. EEG studies on accent effects on cognitive processes 49
2.4. The current dissertation 51
2.4.1. Purpose of this dissertation 51
2.4.2. Overview of the experiments 52
2.4.3. Significance and implications 53
3. Chapter III: An ERP investigation of accented isolated single word processing **53**
3.1. Rationale 53
3.1. Introduction 54
3.2. Methods 58
3.2.1 Participants 58
3.2.2. Materials and Methods 59
3.2.3. Procedure 60
3.2.4. EEG Data Recording and ERP Analyses 60
3.3. Results 62
3.3.1 Behavioral Results 62
3.3.2. ERP Results 62
2.3.2.1. P2 [150-250 ms] 63

3.3.2.2. PMN [250-400 ms] 64
3.3.2.3. N400 [400-600 ms] 65
3.4. Discussion 66
3.5. Conclusion 69
4. Chapter IV: A dual-task investigation of the effects of accented sentence processing on attention and memory 70
4.1. Rationale 70
4.2. Introduction 70
4.2.1. The Present Study 74
4.3. Methods 75
4.3.1 Participants 75
4.3.2 Normative study on audio stimuli for accent strength ratings 75
4.3.3. Procedure and Methods 76
4.3.3.1. Session 1: DPRT task blocks 76
4.3.3.2. Training 77
4.3.3.3. Single Task-Baseline One 78
4.3.3.4a. Dual Task 78
4.3.3.4b. Recordings-Auditory Stimuli of the Dual Task 79
4.3.3.5. Single Task-Baseline Two 79
4.3.3.6. Session 1: Working Memory N-Back Test 80
4.3.3.7. Session 1 and Session 2: Old-New Recognition Memory Test 80
4.3.3.8. Visual Items for Old (Remember, Know) New Memory Test 81
4.3.3.9. Session 2: Questionnaire and Self-Reported Listening Effort 81
4.4. Results 82
4.4.1. Attention 82
4.4.1.1. Listening Effort 82
4.4.1.2. Self-Reported Listening Effort 84
4.4.1.3. Is Listening Effort Related to Participant Background or Working Memory? 85
4.4.1.4. Attentional Results Summary 86
4.4.2. Memory 86
4.4.2.1. Session 1 and Session 2 D' Scores 86
4.4.2.2. Are Memory Scores Related to Participant Background or Working Memory? 87
4.4.3. Attention-Memory Relationship 88
4.5. Discussion 89
4.6. Conclusion 93
5. Chapter V: A frequency investigation of accented discourse processing 95
5.1 Rationale 95
5.2. Introduction 95
5.2.1. The Present Study 99
5.3. Materials & Methods 99
5.3.1 Participants 99
5.3.2. Materials 100
5.3.3. Procedure 101
5.4. Results 102

5.4.1. Behavioral Results 102
5.4.2. Time-Frequency Results 102
5.5. Discussion 106
5.6. Conclusion 107
6. Chapter VI: General Discussion 108
6.1. Summary of findings and general conclusions 108
6.1.1. Linguistic challenges: effect of accent on phonological and lexical-semantic processing in isolated words 108
6.1.2. Cognitive challenges: accent effect on listening effort and memory processing in sentences 110
6.1.3. Cognitive-linguistic integration: effects of accent on extended discourse processing 112
6.2. Broader theoretical implications 113
6.3. Limitations and future directions 116
6.4. Conclusion 117
7. Related publications and manuscripts 119
7.1. Does bilingual status influence automatic speech recognition for young Latino children? 119
7.2. Related publications and conference participation 122
8. References 125
9. Appendices 153

Prelude

Overview of the work

The main objective of my doctoral research project was to explore the effects of speaker characteristics on cognitive processes through both behavioral and electrophysiological evidence. Specifically, I aimed to provide insight into the online and offline cognitive mechanisms of speech processing and whether changes in information retention are due to different attentional demands. Due to the pandemic, the experiments carried out shifted from the original design. While, originally, I hoped to investigate various types of speaker characteristics on attention (accent, face, voice), the focus of the project shifted to an in-depth analysis of accent (foreign, native, dialectal) processing across levels of discourse (single word, sentence, extended speech). The following chapters of this dissertation will present three different experiments that explore how native listeners process different types of accented speech over the time course of three different levels of linguistic analysis. The dissertation will begin with a general overview of how scientists have studied speech processing and how speaker characteristics, specifically accent, have been shown to affect processing. Chapters three through five will contain the main body of the experimental work, addressing how native listeners process non-native and dialectal accent. The final chapter summarizes the findings of the previous three chapters thus drawing the general conclusions of the thesis. A special emphasis will be placed on contextualizing these findings in the broader context of the field to continue to advance the field of accented speech processing research.

1. Chapter I: General Introduction

1.1. Speech processing: background and rationale

I will begin this general introduction by first exploring how speech processing was studied in the past and how technological and scientific innovations have advanced the field. I will then introduce how speech processing can be affected by context-dependent cues and finally will discuss the specific role of speaker accent.

Speech is a form of communication that involves the production and perception of sounds according to the rules of a language. While speech relies on various articulators of the speech apparatus, these movements are coordinated and, in the case of perception, deciphered by the brain. Speech processing, thus, can be broadly described as the successful abstraction of meaning from an acoustic signal. The extraction of information from the signal requires a series of transformations of the acoustic signal as it travels from the auditory perceptual system to the processing areas of the brain. While we describe speech with a variety of linguistic terms, such as phoneme, syllable, word etc, the acoustic signal can only truly be described as meaningful sound produced by particular articulations of the human speech apparatus (Fitch et al., 1997). Speech differs from primary auditory input because of this complex interaction between the auditory system and the language system. Speech can be analyzed on a variety of levels. The phonological level of speech analysis focuses on the sounds of language and how they are used to form words. The lexical-semantic level of speech analysis is concerned with the meaning of words and how sounds combine to form larger units of meaning. Finally, sentence and discourse speech analysis is focused on the structure of sentences and the way that sentences are combined to form conversations. These three levels will be the focus of the present investigation.

1.1.1. Historical perspectives

Scientists have been studying speech processing for centuries, the methodological study of speech production and perception began mostly with phonological observation (see Casserly & Pisoni, 2010) and pathology analysis through patient cases. In the mid-1800s, scientists began to conduct 'deficit' studies with patients or cadavers of patients with neurological disorders as a means to explore the neural underpinnings of language (see Fig.1). Thus, for many years scientists studied the localization of language processes in the brain through deficit, or lesion, analysis. This type of analysis allowed them to learn which aspects of speech are disrupted by the loss of function to a brain area or when stimulating a specific area of the brain. Perhaps the most famous patient cases in the development of this field are those of Dr. Paul Broca and Dr. Carl Wernicke. Broca observed aphasic stroke patients who had lost productive speech abilities. Post-mortum examinations of his famous "Tan" patient, an aphasic patient who was only able to produce the word tan, revealed a lesion in the left inferior frontal gyrus, leading him to theorize that this area, now known as Broca's area, was responsible for productive speech (Broca, 1861). Some 10 years later, another scientist, Carl Wernicke, became well-known for additional aphasic cases. His patients dealt with receptive aphasia rather than expressive aphasia. Wernicke's famous patient, patient "SM", could produce fluent but meaningless speech and had difficulties comprehending spoken or written language. This led to the designation of the left posterior superior temporal gyrus, now called Wernicke's area, as crucially connected to language comprehension (Wernicke, 1874).

Around the 1920s to 1930s, the first recordings of electrical activity to describe speech activities in the brain began. In 1924, a German physicist named Hans Berger became the first scientist to record an electroencephalogram (EEG) of human scalp data. EEG measures voltage fluctuations -on the top of the scalp- that are generated by populations of neurons. It became popular in the late 1930s for its use in clinical neurology. In addition to its clinical applications, scientists still rely on EEG and Event-Related Potentials (ERPs-averaged measurements of a brain response locked to a specific neural event) derived from EEG as a noninvasive technique to study brain activity with high time sensitivity. ERPs of historical significance in speech processing are Auditory Evoked Potentials (AEPs) which are ERP responses to sound stimulation. The study of the auditory evoked potential (AEP) can be traced back to the late 1940s and early 1950s when researchers first began investigating the electrical activity of the brain in response to auditory stimuli as they provide a way to assess the integrity of the auditory system and the brain's response to sound. The AEP remains a valuable tool for the study of auditory processing and the diagnosis of hearing disorders. In addition to the AEP, scientists also began to use techniques like EEG and later MEG to link brain wave frequencies and ERPs with additional specific aspects of speech processing, such as acoustic extraction, normalization between acoustic input and phonological representations, and lexico-semantic integration.

Along with electrophysiological advancements, the mid-20th century saw the rise of the Behaviorism movement of psychology. Behaviorism prioritized the study of observable behaviors, as opposed to internal mental processes like perception and cognition. The movement's influence led to one of the earliest models of speech perception, the motor theory, which will be discussed in the next section. Behaviorism also influenced the field of early speech-processing research, which focused primarily on the acoustic properties of speech and the listener's responses to sounds. One of the key figures in early speech-processing research was B.F. Skinner, a leading behaviorist. Skinner believed that speech sounds were learned through the reinforcement of certain sound patterns and that listeners would learn to associate these patterns with specific meanings over time. Other behaviorists, such as Ivan Pavlov and John Watson, also influenced speech-processing research by emphasizing the importance of stimulus-response associations in learning. While scientists now realize the importance of cognitive mechanisms in speech processing, experiments focusing on behavioral measures such as reaction times are still critical contributors to the growing body of knowledge in the field.

As we get closer to the present day, neuroimaging techniques such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) became widely used for studying speech processing in the 1990s and 2000s. These techniques increased the spatial resolution of the knowledge of the brain, allowing scientists to observe which brain regions were active during different aspects of speech processing and updating previous neurobiological models created based on patient cases.

Finally, it is worth touching on the role of computational modeling in the field. The usage of computational modeling in speech research began around the 1960s and increased around the 1980s and 1990s as advancements in technology and algorithms were made. The Hidden Markov Model (HMM), developed in the 1970s and 1980s, was one of the earliest and most important computer models in speech processing research (see Kayte et al., 2015). HMMs model time-varying spectral vector sequences and are still used as the basis of continuous speech recognition systems. Similarly, neural networks, machine learning algorithms that are used to model various cognitive processes, also were first developed around the 1980s and continue to play an integral role in many fields of research today, including speech recognition and noisy speech processing (see Alshemali & Kalita, 2020).

Although the methods have advanced from early studies, scientists still use a variety of techniques to study speech processing, such as neuroimaging, behavioral experiments, computational modeling, and neural networks. Thanks to these different techniques and their integration several theoretical models on speech processing have been proposed. In the next section, I will describe a few modern models of relevance.

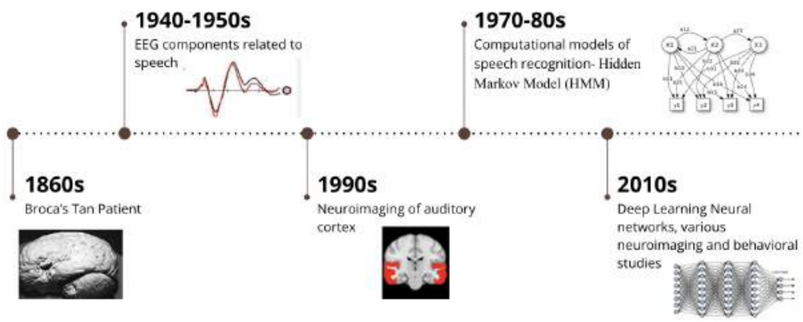


Fig. 1 Timeline of the relevant historical evolution of speech processing studies.

1.1.2. Modern perspectives

Thanks to the combination of the techniques previously described we now have a rather precise picture of the journey of sounds to the brain. Speech processing begins as sound waves are intercepted by the ear. Sound waves enter the ear and cause the eardrum to vibrate, these vibrations continue to the inner ear where the physical signal is converted into an electrical signal. The electrical signal then travels from the cochlea to the auditory nerve and the brainstem where it is processed and filtered. After processing in the brainstem, the auditory signal is sent to the thalamus, which acts as a relay station for sensory information. The thalamus then sends the auditory signal to the primary auditory cortex in the temporal lobe via a bundle of nerve fibers called acoustic radiation. Once the auditory signal reaches the primary auditory cortex, it is further processed and integrated with other sensory information in the brain to give rise to our perception of sound and spread to other brain regions involved in speech processing, such as the superior temporal gyrus, Broca’s area, and Wernicke’s areas (see Fig. 2 for an overview of the primary auditory pathway).

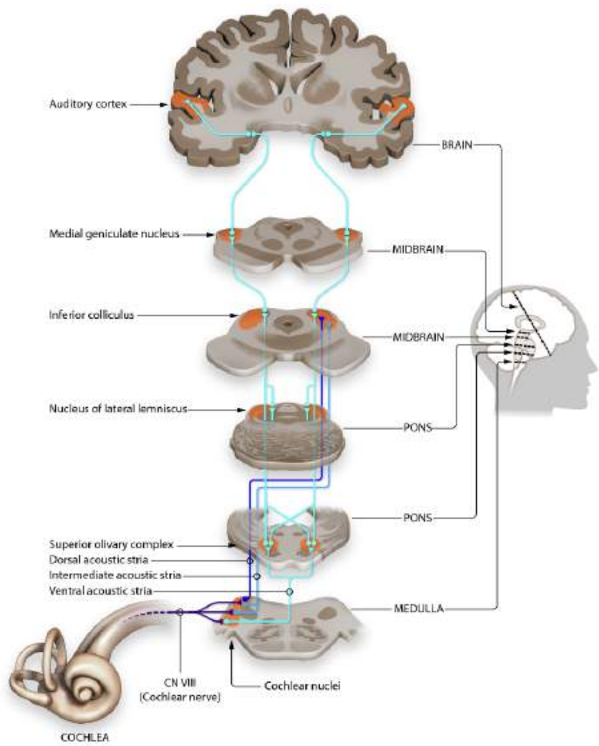


Fig. 2. The auditory pathway. Figure taken from University of Minnesota, Pressbooks Sensation and Perception, chapter on auditory pathways to the brain.

As acoustic information is processed, it is integrated to form a coherent representation of the speech. Linguistic features are combined to form a meaningful interpretation. There are several brain regions important to various aspects of language processing. The following table (Table 1) shows what brain areas can be linked to which linguistic processes.

Brain areas	**Linguistic Processes**
Inferior frontal gyrus (broca's area)	Language production, syntactic analysis
Temporal lobe (wernicke's area)	Language comprehension and semantic analysis
Nerve fibers connecting BA and WA (arcuate fasciculus)	Phonetic information sent between BA and WA
Heschl's gyrus	Acoustic analysis
Angular gyrus/ Superior Temporal sulcus	Converts between visual and auditory stimuli, classify info, abstract thinking

Motor cortex	Directs movement of muscles for articulation
Visual cortex	Visual letter perception
Supramarginal gyrus	Phonological processing, visual word recognition

Table 1. Summary of some relevant language-related brain areas, table taken from Thomas, T., Martin, C. D., Pesciarelli, F., Caffarra, S. (forthcoming). Analyzing Language using Brain Imaging. In P. Gygax and S. Zufferey (Eds.): The Routledge Handbook of Experimental Linguistics. Routledge.

1.1.3. Cognitive and Neurobiological Models of Speech Perception

Over the years, various models have been proposed to explain how the brain processes speech. We know now that multiple areas of the cortex play important functions in speech and language processing (see Fig. 3). Creating a cohesive understanding of how these areas of the brain work together to process speech has been the purpose of neurobiological and cognitive models of speech perception. This section provides a summary of both the current prominent modern neurobiological models of speech perception and also two earlier models. Please note that not all prominent models will be addressed. I will focus on the neurobiological and cognitive models most relevant for the work of the present dissertation, but see Chen & Chang, 2022 for more information.

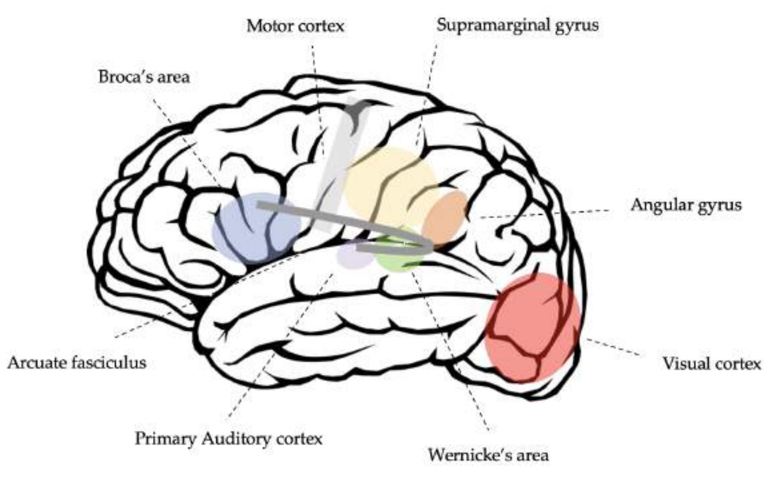


Fig. 3. Schematic of language-related areas in the brain. Figure taken from Thomas, T., Martin, C. D., Pesciarelli, F., Caffarra, S. (forthcoming). Analyzing Language using Brain Imaging. In P. Gygax and S. Zufferey (Eds.): The Routledge Handbook of Experimental Linguistics. Routledge.

Early Models of Speech Perception

Early models of speech perception aimed to understand how the brain processes and comprehends speech. Two contrasting models in this field were the Direct Realist Approach and the Motor Theory.

The Direct Realist Approach (Fowler, 1986) posits that speech perception involves the direct processing of acoustic signals in the auditory system without the need for intermediate representations. According to this model, the brain processes the auditory signals in a bottom-up manner, extracting the acoustic features of speech sounds, such as pitch, intensity, and formants, and directly maps them onto phonetic categories. The Direct Realist Approach emphasizes the importance of sensory information in speech perception and suggests that the brain directly perceives the physical properties of speech sounds.

In contrast, the Motor Theory of speech perception (Liberman & Mattingly, 1985) proposes that perception and production of speech are closely intertwined, with the brain using motor representations to process speech sounds. According to this model, the perception of speech sounds involves the activation of motor representations in the brain that are associated with the articulatory movements used to produce those sounds. This suggests that the brain relies on internal motor simulations to decode speech sounds, linking perception and production in a bidirectional manner.

Modern Neurobiological and Cognitive Models of Speech Perception

In recent years, several models have been proposed to explain the underlying mechanisms of speech perception. The dual-stream model and the TRACE model represent prominent frameworks for understanding speech perception from a neurobiological perspective. While the dual-stream model emphasizes distinct processing streams and contextual integration, the TRACE model focuses on the interaction of phonemic, lexical, and semantic representations. These models provide valuable insights into the neural mechanisms and cognitive processes involved in speech perception, contributing to a deeper understanding of how humans perceive and comprehend speech.

The Dual Stream Model of speech perception (Hickok & Poeppel, 2016) posits that there are two parallel neural processing streams involved in speech perception: a dorsal stream and a ventral stream. The dorsal stream is responsible for processing the sensorimotor aspects of speech, including the mapping of speech sounds onto motor representations, while the ventral stream is responsible for processing the auditory and linguistic aspects of speech, including the extraction of semantic and syntactic information. According to this model, different neural pathways are involved in different aspects of speech perception, with the dorsal stream supporting the sensorimotor processing necessary for speech production and the ventral stream supporting the auditory and linguistic processing necessary for speech comprehension (see Fig. 4). This model highlights the importance of bottom-up acoustic analysis and top-down contextual processing in speech perception.

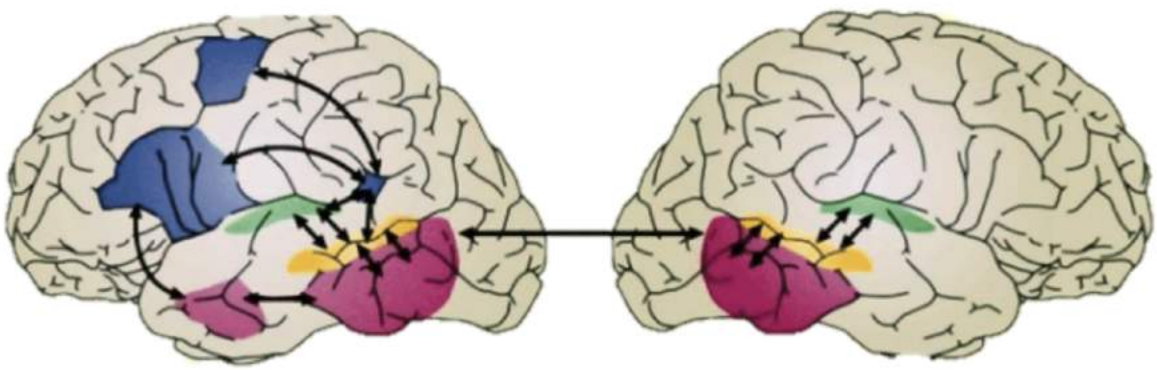


Fig. 4. The Dual-Stream Model, taken from Hickok & Poeppel, 2007.

In contrast, the TRACE model (McClelland & Elman, 1986), takes a connectionist approach, emphasizing the interaction of phonemic, lexical, and semantic representations. This model suggests that speech sounds are processed in a cascading and interactive manner. It describes a network of interconnected nodes where activation flows bidirectionally, with bottom-up processes driven by acoustic input and top-down processes influenced by lexical and semantic cues. TRACE model captures the dynamic competition and interaction between different levels of representation, ultimately shaping the perceived interpretation of speech sounds (see Fig. 5). While varying in their perspectives, both models provide valuable insights into the neural mechanisms and cognitive processes involved in speech perception, contributing to a deeper understanding of how humans perceive and comprehend speech.

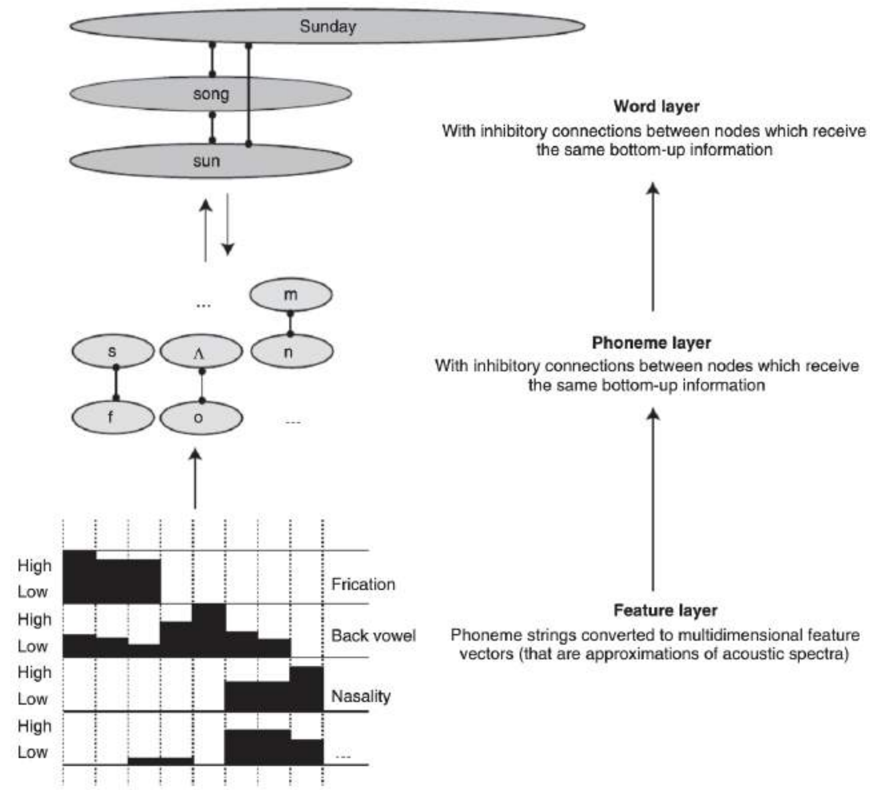


Fig. 5. Recognition process of the word sun by TRACE. Activation in the lower layers flows upwards to the higher levels to all nodes that incorporate the lower layer node. Activation from the word layer also flows back to the phoneme layer. Figure taken from Weber and Scharenborg 2012.

We will revisit these models in the general discussion to situate the dissertation results within a larger theoretical framework. However, before proceeding, it is crucial to acknowledge the significance of context in speech processing. Speech is inherently perceived and processed within a contextual framework, and context plays a pivotal role in shaping and influencing the fundamental aspects of speech processing that have been described thus far.

1.1.3. Role of context in speech processing

An implicit assumption of many early models of speech perception is that due to the variation across different productions, there must be an abstraction of contextual cues in processing that enables the perceiver to map what they hear onto a stored lexical representation. (see Samuel, 2011). However, previous work has demonstrated that contextual information can indeed influence processing (Goldinger, 1996; Nygaard et al., 1994).

Let's imagine a scenario where someone is watching a political debate and one candidate says to the other, "You have a lot of nerve." Without the context of the situation, it would be difficult for the viewer to understand what the politician speaking meant despite having heard the speech clearly. They could mean the comment as an accusation or challenge, or conversely, the same expression could be used to express empathy or admiration depending on what the other candidate had gone through. Thus, it is clear that context matters during speech processing despite not being directly addressed in the aforementioned models. Due to this, additional models accounting for the details of context have been proposed and are discussed below.

1.1.3.1. Context-dependent models and indexical cues

Numerous context-dependent models of speech processing have emphasized the importance of indexical cues in conveying meaning. Indexical cues refer to contextual signals related to the speaker, such as their gestures, facial expressions, tone of voice, and other speaker characteristics that can be used to convey information.

Episodic Model of Lexical Representation

This model (Goldinger, 1998) proposes that listeners store and retrieve detailed episodic traces of specific speech sounds encountered in different contexts, including the indexical cues associated with the speaker's identity. These episodic traces serve as templates for speech perception and influence the listener's perception of speech sounds from the same speaker in the future.

Standard Two-Step Model of Language Interpretation

This model (Hagoort & van Berkum, 2007) suggests that listeners initially process speech sounds in a pre-lexical stage, where they extract basic acoustic-phonetic features from the speech signal. In the second step, listeners integrate the extracted features with higher-level linguistic knowledge, including indexical cues, to map the acoustic-phonetic features onto lexical representations.

Both models highlight the importance of indexical cues in shaping context-dependent speech perception. The episodic model emphasizes the role of stored episodic traces, while the standard two-step model emphasizes the integration of indexical cues with other linguistic knowledge in the mapping process. Together, these models provide valuable insights into the complex interplay between acoustic-phonetic information and higher-level linguistic knowledge, including indexical cues, in speech perception. These models are at least partially consistent with the dual-stream neurobiological model of speech processing, which suggests that separate neural pathways are responsible for processing acoustic-phonetic information and integrating it with higher-level linguistic knowledge, including indexical cues.

1.1.3.2. Speaker identity as an indexical cue

One particularly salient indexical cue is speaker identity, which can be reflected by acoustic, vocal, and social-grouping characteristics of speech. Speaker identity is a multifaceted, complex and dynamic cue from which the listener can infer a combination of various factors, including biological, physiological, social, and cultural aspects of his/her interlocutor. Specifically, the speaker's accent can transmit information about the regional, cultural, and linguistic background of the interlocutor and consequently affect how speech is processed. The following sections will define and delve deeper into speaker accent and how this indexical feature affects speech perception and processing on a multitude of levels.

1.2. Accent perception and processing

1.2.1. Accent as a case of speaker identity

Accent is a speech quality that every person has and that is largely determined by the geographic, linguistic, and cultural environment in which we have learned a language. An accent, in its broadest sense, refers to the way in which a person speaks a language. Accents are often associated with particular geographic regions, ethnic or cultural groups, first language, and socio-economic status. This is influenced by various acoustic features such as pronunciation, intonation, prosody, and stress patterns of speech. Aspects of accent also include phonemic, lexical, syntactic or semantic variations related to a specific region or group, such as phonological approximations and grammatical inaccuracies (Flege, 1995; Saito et al., 2016; Franceschina, 2001; Caffarra & Martin, 2019; Mariko, 2007).

1.2.1.1. Accent types

It is important to note that for the purposes of the research carried out in this dissertation, I will be broadly categorizing accents into two larger categories and further subdividing one of those said categories. Namely, I will be using the terms native, foreign, and dialectal to reflect native (native and dialectal[1]) and non-native (foreign) accent variations[2]. Non-native speech is the term used to describe speech produced by individuals who speak a language other than their first language. The distinction between native and non-native languages seems to rely on the age when the language was acquired (see Nikolov & Djigunović, 2006 for a review). When non-native languages are learned later in life, typically from late adolescence to adulthood, the speaker will likely have a non-native accent (Birdsong, 2006). Compared to native speech, non-native speech usually shows deviations across various dimensions, including both segmental (i.e. vowels & consonants) and suprasegmental (i.e. prosody) aspects (Baese-Berk et al., 2020). Both dialectal and native accent have been traditionally described to fall under a ‘native’ accent category. However, a dialectal accent will differ from the native accent in that it is not local to the area of the chosen native accent and thus will contain various speech variations. Despite this, dialectal speech variations differ from foreign speech variations in that they contain native phoneme variations. Foreign accents, conversely, are generally thought to be characterized by an extra amount of acoustic variability (see Vaughn et al., 2019 for a review). They often contain phonemes and syntactic patterns that do not exist in the language that they are speaking as they use their own native language’s phonological system as a filter for both perceiving and producing speech.

In speech research, accents are usually defined according to the experimental group of interest as they are relative terms. For instance, let’s suppose one conducts their research in the

[1] Some studies refer to dialectal accents as regional accents. For the sake of consistency, I will use the term dialectal.

[2] It should be further noted that some scholars have proposed alternate terminology to reflect the diversity of language profiles and experiences that are not represented fully by a native/non-native binary categorization (for further discussion see (Cheng et al., 2021). However, to stay in line with the majority of the literature these are the terms I will be using at this time.

northern region of Cantabria, Spain. If one wants to study how a local Cantabrian group of Spanish native speakers process different accents, they call the accent of this local group a native accent and define the others accordingly. A dialectal accent for this group could be a Spanish native speaker from a country such as Mexico or Colombia, characterized by more aspirated sounds (Marden, 1896), or could be a Spanish native speaker from another region of Spain with a distinct accent, such as Andalusia. A foreign accent for this group would be someone with a native language other than Spanish but who has learned the spoken Spanish language (often later in life). For instance, this could be someone from England or China speaking Spanish. In the case of Chinese-accented Spanish, there are substitutions of trills and taps with liquids (non-native phonological variations) (Mateu, 1996). Whereas with English-accented Spanish, there are usually more grammatical gender errors (non-native grammatical variations) (Franceschina, 2005). The speech of foreign-accented Spanish is in many ways less reliable than either of the native accent category variations. Errors are often more inconsistent, and often there is more individual variation in the accent strength of individual speakers (Witteman et al., 2014). This means that foreign-accented speech presents the listener with a variety of challenges that other accents do not. The research questions explored throughout this dissertation are approached from the listeners' perspective. Thus, all presented studies consider how a native listener processes speech from native and non-native speakers and as such it is necessary to explore the challenges of non-native speech accent. Non-native accent challenges posed to the native listener are not only linguistic but also more broadly cognitive. In the next section, I will provide an overview of some of the main challenges.

1.2.1.2. Challenges of foreign-accented speech

Linguistic challenges

Phonological challenges

Foreign-accented speech often consists of acoustic realizations that differ from the listener's representations of the sounds of their language. Many linguistic sounds are not shared between languages and varying proficiency levels of non-native speakers lead to a spectrum of acoustic productions of sounds that do not exist in the native language of the speaker. Even when the target sound exists in the native language of the foreign-accented speaker, their knowledge of their own native orthography and its associated sound may interfere with their pronunciation of the target sound. For instance, American English and Spanish share the tap /ɾ/ but orthographically the sounds do not overlap. While the Spanish word pero contains the /ɾ/ for the sound that corresponds to the orthographic r, in English, taps do not function as rhotics but are realizations of intervocalic apical stops, as in words like butter or city. Because of this incongruency, an English native L2 speaker of Spanish is much more likely to produce the rhotic /ɹ/ in this situation because that is the sound most often associated with the orthographic r. This means that to effectively recognize spoken words with non-native accents, listeners must alter their perception of these relatively novel sounds. Chapter three will focus on how native listeners deal with these types of challenges.

Lexical-semantic challenges

Foreign-accented speech also introduces difficulties with combining words to deduce the intended meaning. Foreign-accented speakers are more likely to incorrectly select lexical items. They may substitute based on phonological or semantic similarity to the target word, or they make incorrect direct translations from their native language. They are also likely to make grammatical errors or incoherencies. Because of this, the listener must be able to access a variety of potential meanings, hypothesize multiple interpretations and select the intended representation for any given foreign-accented sentence. Chapters three and five will consider the challenges of lexical access in accented speech listening.

Cognitive challenges

Attentional challenges

While the empirical evidence on this topic is less developed, foreign-accented speech has been thought to pose a greater cognitive load on the attentional system of the listener. Cognitive load refers to the cognitive demands placed on the attentional and working memory system during a cognitive task (Pichora-Fuller et al., 2016), while cognitive effort refers to the allocation of cognitive resources for task performance (de Jong, 2010). Foreign accent is associated with greater cognitive load because the variations in pronunciation, intonation, or rhythm in accented speech may capture listeners' attention and draw their focus away from other linguistic cues, such as the semantic and syntactic aspects of speech. As a result, listeners may need to exert more cognitive effort to selectively attend to relevant linguistic cues in accented speech, which can increase the cognitive load during speech perception. Furthermore, listeners may also face challenges in shifting their attention between different sources of information in accented speech, such as the speech sounds and the speaker's accent. For instance, listeners may need to constantly switch their attention between processing the unfamiliar accent and deciphering the intended message, which can disrupt the smooth processing of speech and impact comprehension. This challenge will be the focus of chapter four.

Memory challenges

There has been limited evidence to suggest that foreign-accented speech affects the level of encoding of memory representations of listeners. The variations in pronunciation, intonation, or rhythm in accented speech may make it more difficult for listeners to encode and store the speech sounds and lexical items in their memory. This is because the deviations from the standard pronunciation or prosody may not conform to listeners' existing mental representations of the language or dialect, and as a result, may not be as readily encoded and retrieved from memory. Moreover, accented speech may also challenge listeners' working memory capacity, which refers to the cognitive system responsible for temporarily storing and manipulating information during complex tasks, such as speech perception. The additional cognitive load incurred by processing accented speech, including the need to accommodate the unfamiliar accent, may tax listeners' working memory capacity, leading to reduced processing resources for other cognitive processes involved in speech perception, such as comprehension and integration of meaning. In addition to an effect on working memory demands and memory encoding, foreign accents may also impact long-term memory by affecting information retention periods. This challenge will be further explored in chapter four.

Cognitive-linguistic integration challenges

Listening to extended discourse in a foreign accent poses unique challenges for listeners as it requires integrating both linguistic and cognitive processes. Listeners must process the incoming speech sounds and map them onto the phonological and lexical-semantic representations in their mental lexicon. They must also engage in higher-level cognitive processes such as attention, working memory, and discourse comprehension to derive meaning from the speech. Integrating these linguistic and cognitive processes requires listeners to allocate cognitive resources efficiently, manage competing information, and make inferences based on contextual cues. Consequently, listening to extended discourse in a foreign accent can be a challenging task that requires both linguistic and cognitive flexibility. This challenge will be addressed in chapter five.

1.2.1.3. Competing theoretical frameworks of accent perception and processing

Many studies have shown that non-linguistic aspects of speech related to interlocutor identity, including accent, are retained and affect processing (Goldinger, 1996; Bradlow et al., 1997). Indeed, the previously mentioned challenges have been shown to have lasting effects on processing. Thus, to address these challenges various theoretical frameworks have been proposed. These frameworks differ from the context-dependent models described above in that they are specific to the challenges of non-native accent processing as opposed to the broader integration of several context-dependent cues. Below some of the theoretical frameworks for accent processing are discussed in the order of the challenges previously presented.

Linguistic challenges

The Perceptual Assimilation Model (PAM), proposed by Best (Best & Others, 1994), and the Speech Learning Model (SLM), proposed by Flege (Flege, 1995), provide perspectives on the perception of non-native speech sounds and the role of linguistic experience.

PAM emphasizes the role of native language categories in shaping the perception of non-native speech sounds, suggesting that when encountering non-native speech sounds, listeners assimilate them into their native language sound categories. This assimilation process involves comparing the non-native sounds to native language sounds and categorizing them based on the closest match. This assimilation process may result in non-native speech sounds being perceived as similar to native sounds, or as distinct from any native category. PAM also suggests that the degree of similarity between non-native and native sounds influences their perceptual assimilation, with more similarity leading to better assimilation.

In contrast, SLM focuses on the role of speech learning and experience in accented speech perception. According to SLM, listeners' previous experience with non-native speech sounds, such as exposure to a second language, influences their perception of accented speech. SLM proposes that listeners develop new categories for non-native speech sounds based on their prior experience. These newly formed categories may not necessarily align with the native language sound categories, resulting in a distinct perception of accented speech sounds compared to native speech sounds. SLM highlights the influence of individual speech learning experiences on shaping the perception of non-native sounds. These models offer different perspectives on how listeners perceive and categorize non-native speech sounds, showcasing the importance of both native language influence and individual speech learning experiences in shaping the perception of unfamiliar sounds.

Another theoretical perspective of relevance is that of Cascading Effects. As has been previously established, foreign accent perception can be understood as a series of physical acoustic deviations from "standard" native phonological categories. Consequently, native listeners processing non-native speech must contend with the persistent disparity between their own underlying phonological representations and the foreign-accented input they perceive. These phonological challenges are believed to hinder single-word recognition, which in turn may have repercussions on higher levels of language comprehension, including semantics and syntax. It has been proposed that the pre-lexical phase of spoken word recognition involves a normalization process aimed at filtering out speaker-specific acoustic divergences, including accents of the speaker (Goslin et al., 2012). In the context of foreign-accented speech processing, the normalization process strives to minimize the gap between the underlying phonological representations of native listeners and the acoustic distortions produced by non-native speakers (Goslin et al., 2012; Sumner & Tilsen, 2011). While empirical evidence on this topic is still limited, it appears that the extent to which normalization occurs varies depending on the sentential context and the type of accent. For instance, listeners seem to be more successful in normalizing familiar or native-like accents, particularly in the case of single-word recognition (Goslin et al., 2012). However, when it comes to the comprehension of complete utterances, most studies indicate that foreign-accented sentence processing is accompanied by minimal or even absent normalization. As a result, accent-related processing issues cascade into later stages of word analysis and sentence comprehension, specifically in the domains of semantics and syntax (Goslin et al., 2012; Gosselin et al., 2021). Consequently, in most natural conversational contexts (where conversations typically involve complete sentences rather than isolated words), foreign-accented speech characteristics cannot be easily filtered out during an early phase of phonological analysis, and cascading effects are detected within higher-order cognitive processes.

While the previously mentioned models of accented processing focus on the phonological challenges of native vs non-native speech processing, there are also specific theories that focus on how to categorize the acoustic variations presented by dialectal accented speech. Two competing hypotheses in this domain have been proposed. The Different Processes Hypothesis postulates that there are qualitative differences in the processing mechanisms recruited for dialectal and foreign speech normalization (Floccia et al., 2009). The Different Processes Hypothesis posits that there is a categorical difference between foreign and native accent variations, meaning that there is something uniquely challenging about non-native speech variations, regardless if the native variation is a local accent or a dialectal accent.

The Perceptual Distance Hypothesis states that the mechanisms underlying dialectal accent processing are attenuated versions of those of foreign (Clarke & Garrett, 2004). The Perceptual Distance Hypothesis relies upon the knowledge that accents can be placed on a perceptual scale based on their acoustic distance from native speech, with foreign accents being the furthest from native and dialectal accents falling somewhere in between. According to this hypothesis, dialectal accent processing mechanisms would be attenuated versions of foreign-accented processing mechanisms. These theories will be disambiguated in chapters three and five.

Attentional and memory challenges

The processing of accented speech also poses significant attentional challenges. Three theories—Cognitive Load Theory, Cognitive Effort Theory, and Indexical-Free Theory—offer distinct perspectives, elucidating the allocation of attentional resources and its potential subsequent impact on task performance and memory retention. The first two of these theories operate under the assumption of Exemplar models while the third is not contained in the Exemplar based theories.

Cognitive Load Theory (see Paas, Tuovinen, et al., 2003) posits that cognitive resources are inherently limited, and tasks with higher cognitive load tend to impede performance and hinder information retention. As the effort required for a task increases, the cognitive load of the task also increases. Within this framework, the comprehension of foreign-accented speech necessitates additional cognitive resources, resulting in more difficulties in processing foreign-accented speech than native-accented speech. Thus, in addition to its application in fields such as education and memory, Cognitive Load Theory has been used to explain decreases in performance during foreign-accented speech comprehension and retention.

Conversely, the Cognitive Effort Theory (see Mayer et al., 2005; Tyler et al., 1979) argues that increased cognitive effort associated with higher cognitive load may actually facilitate task performance and improve information retention. According to this perspective, the increase in cognitive load of foreign-accented listening triggers a reallocation or re-prioritization of cognitive resources, thereby enhancing attention directed toward the task at hand. Consequently, the increased cognitive load involved in processing foreign-accented speech has the potential to bolster retention and overall task performance through the reallocation of attentional resources driven by cognitive effort.

The effect of foreign-accented speech on memory is something under continued investigation and is studied under the framework of exemplar models. Exemplar models of speech perception posit that utterances are stored in the mind as distinct exemplars, retaining detailed acoustic and phonetic information (Foulkes & Docherty, 2006; Goldinger, 1998). Social information pertaining to the speaker is also encoded during perception and linked to the exemplar in memory (Foulkes & Docherty, 2006; Hay et al., 2006). As such, the cognitive representation of an utterance is not an abstract underlying form, but rather a distribution of stored exemplars. Speech perception entails mapping the acoustic signal to the stored distribution based on similarity to both the acoustic and social information associated with the exemplar (Johnson, 1997). This hypothesis is corroborated by studies demonstrating that social factors such as the speaker's dialect, gender, age, and social class heavily influence speech perception (Nygaard & Pisoni, 1998; Bradlow et al., 1999; Johnson et al., 1999; Hay et al., 2006). Consequently, many researchers now consider variability as a fundamental aspect of language comprehension and advocate for the inclusion of variability in speech perception models (Luce & Mclennan, 2005); Nygaard & Pisoni, 1995). Chapter four will investigate the effect of foreign accent on attention and memory.

In contrast, the Indexical-Free Theory posits that accent, as an indexical property of speech, should not impact memory performances (Romero-Rivas et al., 2015). Drawing upon the models presented in section 1.1.3, wherein the contextual information is abstracted during speech processing, this theory contends that indexical properties are detached prior to memory encoding or retrieval, thereby exerting no influence on these cognitive processes.

These three theories present competing viewpoints regarding the allocation of attentional resources and the subsequent memory consequences of processing accented speech. Consequently, further research is warranted to comprehensively grasp the intricate interplay between cognitive load, cognitive effort, and the processing of accented speech, as well as their implications for performance and memory retention within diverse contextual frameworks. Chapter four will investigate the effect of foreign accent on attention and memory in the context of these theories.

2. Chapter II: Experimental evidence on accent perception and processing

Traditionally, theoretical frameworks in the field of speech perception have been tested through behavioral means, which predominantly employ response time and accuracy measures. While these studies provide valuable offline information regarding speech processing, they do not allow scientists to access online processing. Because speech is a medium that unfolds very rapidly, information at a highly sensitive temporal scale is useful to further understand what is happening during perception and processing as opposed to solely the behavioral effects of processing. Because of this, the studies presented in chapters three and five will adopt the highly time-sensitive technique of Electroencephalography (EEG) in addition to collecting behavioral measures. Below the methodology of EEG will be presented, followed by behavioral and electrophysiological findings on accent processing.

2.1. Methodology

2.1.1. EEG

Electroencephalography is a non-invasive neuroimaging tool that measures neuronal activity at the scalp. It is an especially useful tool for the study of fine-grain time-sensitive information for language processes in the brain. EEG is recorded with multiple electrodes that are placed on the scalp of a participant and are able to track fluctuations in electrophysiological brain activity. This technique is used to map brain activity by measuring the electrical fields generated by extracellular currents of brain cells (Berger, 1929). However, it can only detect the synchronized post-synaptic activity of thousands of pyramidal neurons acting in synchrony (Hämäläinen & Hari, 2002). With EEG, researchers can isolate and temporally localize the neural correlates of a variety of specific language-related processes on a millisecond scale. Fine-grained temporal resolution enables researchers to isolate the processes that are related to the analysis of specific linguistic information presented at a specific time point, such as the time course of word processing within a sentence. The electrophysiological signal obtained from EEG is multidimensional, i.e., it can be described with time, space, amplitude, polarity and/or frequency of brain activity. This multidimensionality means that different analyses can be performed on distinct dimensions of the data generated from this technique, obtaining different types of EEG measures.

2.1.1.1. Evoked activity time domain

Time domain analyses, such as Event-Related Potentials (ERP), are one set of popular EEG measures. ERPs are the averaged measures of extremely small voltages produced when a large number of neurons fire simultaneously in response to a single internal or external event (e.g. an ocular movement or the presentation of a spoken word during an experiment, respectively). They enable researchers to quantify the psychophysiological correlates of the mental operations connected to the presentation of a particular event. ERPs are characterized by topographical distribution (where they peak on the scalp) and polarity (positive or negative) and timing (when they occur). These characteristics can be used to distinguish ERP components linked with a particular functional interpretation of sensory or cognitive processes. ERPs consist of positive and negative deflections known as components, each reflecting a different cognitive process or brain activity related to the stimulus. Common ERP components include the P1, N1, P2, N2, and P3, among others (Sur & Sinha, 2009; Luck, 2014). ERPs provide valuable insights into various cognitive processes, such as perception, attention, memory, and language processing. They are widely used in neuroscience and psychology research to study how the brain responds to different stimuli and how cognitive processes unfold over time. In the context of the dual-stream model, ERPs have been utilized to investigate the distinct processing streams and their functional roles. Specifically, researchers have examined the P2 component of ERP, which is associated with early sensory processing and the discrimination of speech sounds. They have also examined the PMN, and N400 in relation to phonological mapping and lexical integration. The dual-stream model suggests that both streams, responsible for processing acoustic properties in different ways, play a role in generating these ERP components (see Lau et al., 2008; Brouwer & Hoeks, 2013).

Table 2 below provides an overview of components related to language processes, which are highly relevant for the study of speech accent and will be considered in chapters three and five.

Component	Time window	Distribution	Related processes
P2	150-250	Centro-Frontal	Extraction of acoustic features in speech (Reinke et al., 2003) It is believed to reflect early sensory processing and is often associated with the detection and categorization of auditory stimuli, including speech sounds (Reinke et al., 2003). It has been observed to vary in amplitude and latency depending on factors such as stimulus characteristics, attention, and individual differences
PMN	250-300	Fronto-central	Sound analysis It has been related to normalization between acoustic input and phonological representations (Newman & Connolly, 2009). Phonological incongruencies, and/or a mismatch between perceived speech input and a lexical representation, are thought to elicit a larger PMN (Connolly et al., 2001; Connolly & Phillips, 1994).
N400	300-500	Centro-posterior	Semantic analysis It has been related to aspects of lexical-semantic analysis and was initially observed in response to outright semantic violations (e.g., “He spread the warm bread with socks” as compared to “He spread the warm bread with jam”). The N400 effect typically consists of a greater negative amplitude for the violated as compared to the control condition (Kutas & Federmeier, 2011).

Table 2. Some electrophysiological correlates of language processing, table adapted from Thomas, T., Martin, C. D., Pesciarelli, F., Caffarra, S. (forthcoming). Analyzing Language using Brain Imaging. In P. Gygax and S. Zufferey (Eds.): The Routledge Handbook of Experimental Linguistics. Routledge.
.

2.1.1.2. Oscillatory activity frequency domain

Another EEG analysis consists of time-frequency decomposition. With this method, the oscillatory pattern of brain activity is also taken into consideration. The decomposition of the brain signal using mathematical techniques based on Fourier analysis yields oscillatory activity indicators. Time-frequency analysis is a multivariate approach that enables the investigation of parallel cognitive processes across several frequency scales yet occurring within the same temporal window. This approach enables the decomposition of the brain signal into many components depending on various frequency ranges, i.e., neural oscillations, as opposed to reducing the brain signal to a single variable. The frequency bands are typically divided into five categories: delta (1-4 Hz), theta (4-8 Hz), alpha (9-12 Hz), beta (13-30 Hz), and gamma (> 30 Hz). The benefit of using this technique is that it allows for the separation of frequency bands, which can be correlated with distinct linguistic and acoustic processes. However, this advantage comes at the cost of losing detailed temporal information provided by the ERP technique. Due to this, the two techniques are complementary, enabling the study of both top-down and bottom-up mechanisms. Within the context of the Dual-Stream Model, gamma band and theta band oscillations have been related to parallel routes in the mapping from acoustic input to lexical phonological representations, where gamma is related to segment level representations and theta to syllable level[3] (Hickok & Poeppel, 2007).

2.2. Behavioral studies on accent effects on cognitive processes

The challenges discussed in section 1.2.1.2. have provided the catalyst for a multitude of behavioral studies on how the native listener processes non-native speech. The majority of behavioral studies have collected accuracy and reaction time measures in order to estimate the difficulties listeners encounter processing foreign-accented speech at the linguistic, attentional and memory level.

Linguistic challenges

Many researchers agree that variability, such as foreign accent, is a fundamental aspect of language comprehension (Luce & Mclennan, 2005.; Nygaard & Pisoni, 1995; Cristia et al., 2012). Indeed, non-native speech tends to be slower than native speech (Guion et al., 2000); (Mok & Dellwo, 2008) and word duration of foreign-accented speech is generally longer than native speech (e.g., Romero-Rivas et al., 2015), which may result in a later appearance of the point of auditory uniqueness (i.e., the first point in the acoustic signal where a unique lexical candidate can be recognized) thus contributing to listeners' difficulties in foreign-accented

[3] see Fig. 27 in discussion.

speech processing (Romero-Rivas et al., 2022). Listeners are less accurate when transcribing the speech of both foreign-accented speakers (Gass & Varonis, 1984) and dialectal-accented speakers (Mason, 1946; Labov et al., 1997). Additionally, the intelligibility of both foreign-accented speech (Rogers et al., 2004) and dialectal-accented speech (Clopper & Bradlow, 2008) is more affected by background noise than native speech and is generally rated as lower compared to native speech. As a result of this, native listeners are usually slower/less accurate at phoneme detection, speech transcription and speech comprehension tasks, leading listeners to rely more on lexical information (Bürki-Cohen et al., 2001).

Several studies have also found that listeners' lexical-semantic processing is affected by foreign accent with listeners' performance dropping at word recognition, lexical retrieval, truth evaluation, and semantic categorization tasks (Lane, 1963; Schmid & Yeni-Komshian, 1999); (van Wijngaarden, 2001). Foreign-accented words and sentences are more frequently misidentified than native-accented ones(Lane, 1963; (Braun et al., 2011; Clarke & Garrett, 2004; van Wijngaarden, 2001; Munro & Derwing, 1995b; Anderson-Hsieh & Koehler, 1988; Schmid & Yeni-Komshian, 1999). Foreign accent has been shown to have a greater impairment on lexical retrieval (Girard et al., 2008; Floccia et al., 2009), with slower word recognition observed for words uttered in an unfamiliar dialectal accent, and even more slowly in a foreign accent (Floccia et al., 2006). Studies have shown that when performing semantic tasks, such as truth evaluation judgments or semantic categorization tasks, foreign-accented speech results in longer reaction times (Adank, 2012; Adank & Janse, 2009; Munro & Derwing, 1995b; Adank & McQueen, 2007). However, despite this longer processing time, listeners are generally able to achieve similar levels of comprehension accuracy for both native and non-native accents (Grey & van Hell, 2017; Foucart & Hartsuiker, 2021). Therefore, even with foreign-accented speech, listeners are still able to accurately comprehend the semantic content, despite the additional processing time required.

Some behavioral evidence has been observed for the PAM model discussed above. By finding that native listeners attempt to 'normalize' unfamiliar speech sounds, the idea of assimilation of sounds into native categories is supported. While behavioral evidence for both hypotheses differentiating dialectal and foreign accent has been found. Processing speed and word recognition rates have both been shown to be more impaired by foreign than dialectal accents, though impaired by both compared to native which supports the Perceptual Distance

Hypothesis (Adank et al., 2009; Floccia et al., 2006). However, other studies have found that dialectal accent is often treated like native accent while foreign accent is processed differently, supporting the Different Processes Hypothesis (see Floccia et al., 2009).

Cognitive Challenges

Attentional challenges

It has been shown that listeners recruit additional cognitive resources in adverse listening conditions (Heinrich et al., 2008; Rabbitt, 1968). Previous studies have demonstrated that foreign-accented speech may be more difficult to attend to than native-accented speech (Brown et al., 2020a; Lev-Ari & Peperkamp, 2014; McLaughlin & Van Engen, 2020; Munro & Derwing, 1995b; Van Engen & Peelle, 2014). Studies have shown that native English-speaking individuals make fewer transcription errors when listening to native English speakers in comparison to speakers with accents from languages such as Korean or Chinese (see Bent et al., 2003). Foreign accented speakers are perceived as less comprehensible compared to native speakers and are associated with a slower processing speed, often thought to be indicative of cognitive effort (Munro & Derwing, 1995b). Most behavioral evidence linking the increase in cognitive effort associated with foreign accent processing to task performance aligns with the Cognitive Load Theory (e.g., reduced accuracy/longer response times for foreign-accented compared to native-accented speech), rather than the Cognitive Effort Theory.

Memory challenges

Research on memory effects of foreign-accented speech is more limited and varied. Research of accent processing has found that listeners have less detailed representations of the speech of non-native speakers (Lev-Ari & Keysar, 2012; 2017), while dialectal accented speech does not seem to impact the memory or credibility of the message (Frances et al., 2018). One study showed that short-term memory performance decreases when listening to either a computer-generated voice or a foreign-accented speaker, as compared to a native-accented speaker (Atkinson et al., 2005). Further supporting this claim, another study found that the presence of foreign accent in a crime perpetrator impaired witnesses' memory of the physical appearance of the perpetrator (Pickel & Staller, 2012). Conversely, a recent study by Foucart & Hartsuiker found no differences in a sentence memory task between native and foreign-accented trials (Foucart & Hartsuiker, 2021). While another study by Cho & Feldman found beneficial memory effects of foreign-accented utterances. They found that English words spoken in an unfamiliar Dutch accent were more likely to be recalled and recognized than words in a more familiar accent (Cho & Feldman, 2013). Based on the mixed results of these studies and assuming that foreign accent is associated with higher cognitive effort than native accent, support for both the Cognitive Load Theory (e.g. impaired memory performance) and the Cognitive Effort Theory (e.g. enhanced memory performance) have been found in memory. However, generally, there has been more support for the Cognitive Load theory, with more studies showing impaired memory for foreign accents.

2.3. EEG studies on accent effects on cognitive processes

This section examines the empirical evidence from EEG studies investigating the impact of accents on cognitive processes, with a particular focus on linguistic challenges, attentional challenges, and memory challenges.

Linguistic challenges

EEG studies have revealed distinct neural responses associated with the perception and comprehension of accented speech. Studies investigating event-related potentials (ERPs) have supported that foreign-accented speech is associated with increased difficulty in extracting acoustic features like phonemes, showing a reduced P200 component for foreign-accented speech compared to native (Romero-Rivas et al., 2015; Foucart et al., 2020). Another early component affected by accent is the Phonological Mapping Negativity (PMN). One study found that when participants listened to sentences spoken with an unfamiliar dialectal accent, the PMN response evoked by the final word was stronger compared to sentences produced in their own accent. However, the PMN response was not observed for speech with a foreign accent (but see Porretta et al., 2017). Some studies have reported inconsistencies in the patterns of semantic processing. Semantic violations in foreign-accented speech have been associated with delayed N400 responses compared to native (Grey & van Hell, 2017), late negativities (Grey et al., 2019), similar N400 responses (Hanulíková et al., 2012), or an N400 followed by a late positivity in the native condition (Romero-Rivas et al., 2015). Grey and Van Hell theorized that the delay of the N400 in foreign-accented sentences suggested slower lexical access for foreign-accented speech (Grey & van Hell, 2017; Grey et al., 2020). Jiang et al., 2020, found that doubtful tone generates N400 and late negativity for foreign and dialectal accented speech but not native. Additionally, predictions of upcoming words during sentence comprehension are less efficient when listening to foreign-accented speech (Romero-Rivas et al., 2016). Taken as a whole, the ERP evidence of accent processing shares some degree of alignment with the Cascading Effects model of accent processing. There are cascading effects of phonological analysis on higher levels of foreign-accented speech processing: the modulation of the N400 by accent indicates that the level of phonology can cascade into the semantic interface. Furthermore, ERP evidence for both the Perceptual Distance (see Jiang et al., 2020) and the Different Processes hypothesis (see Goslin et al., 2012) can be observed.

Cognitive Challenges

EEG studies have also attempted to shed light on the attentional mechanisms involved in accent processing. For example, more pronounced N2 amplitudes have been interpreted as reflecting increased cognitive processing in foreign-accented speech processing (Nieuwenhuis et al., 2004; Dickter et al., 2012). The N2 component, occurring around 200-300 ms after stimulus onset, has exhibited increased amplitudes when participants allocate attentional resources to the perception of non-native accented speech (Gow, 2012). This suggests that individuals engage in heightened attentional processing to decipher and interpret phonetic information embedded in accented speech. One study found that the ERP component N2 was more negative for incongruent accent-face pairings, suggesting that incongruence was related to more effortful processing of accented speakers (Hansen et al., 2017). Another study by Hatzidaki & colleagues found foreign-accented speech interfered with the processing of positive valence instead favoring a negativity bias in processing emotionally salient words, which they believe suggested heightened attention to negative words with foreign-accented speech (Hatzidaki et al., 2015). These findings imply that accents can trigger an enhanced cognitive engagement to facilitate accurate comprehension, which if performance remains high could support the Cognitive Effort Theory. As mentioned in the previous section on behavioral evidence, EEG evidence on the effect of accent on memory is limited. In-group membership has been associated with enhanced memory retrieval (Wolff et al., 2014; Wiese, 2012; Wilson & Hugenberg, 2010). One study investigated the in-group memory advantage with native and foreign languages and found that faces were recalled better if they shared a native language with the listener both behaviorally and in the ERP analysis (Baus et al., 2017). This finding would be, conversely, in line with the Cognitive Load Theory due to the decrease in performance. Distinguishing between the Cognitive Effort and Cognitive Load theories in ERP studies presents challenges due to inconsistent findings that can be open to interpretation. For instance, an enhanced N400 response could suggest that semantic processing was demanding, whereas the absence of an N400 might indicate that processing was excessively demanding, leaving the listener with insufficient cognitive resources to complete the semantic analysis (Cognitive Load Theory). However, in cases where the semantic manipulation requires fewer resources, accent effects may emerge that favor accented speech (Cognitive Effort Theory). While additional studies are necessary the combined evidence from behavioral and EEG studies suggests there is at least some truth to the Cognitive Load framework.

2.4. The current dissertation

2.4.1. Purpose of this dissertation

Throughout these first two chapters, I have reviewed the experimental and theoretical literature underlying speech processing and perception with a specific focus on the role of the speaker's accent. While many previous studies have shown that foreign accent affects speech processing and fewer have suggested that it may have residual effects on complex processes such as memory representations, there is less of a consensus on to what extent and how those effects change throughout the time course of linguistic processing. There is furthermore still an ongoing debate about how dialectal processing varies in relation to both native and foreign accent. In an earlier section (section 1.2.1.3), I introduced two hypotheses of dialectal accent processing. Previous studies have contributed to the plausibility of both previously introduced hypotheses, and importantly for our purposes, previous electroencephalography experiments exploring the question have mainly used sentences as material. Further research is necessary to explore whether foreign accents are processed uniquely compared to all forms of native speech, including both native accent and dialectal variations. It remains to be determined if dialectal accent is perceived differently from native accent, despite both being variations of native speech. Based on the models and evidence described above we know that there are contrastive results on the effect of different types of accent on the time course of language analysis and there are at least three main research questions that remain unsolved.

Accordingly, the central aim of this dissertation remains to further investigate processing mechanisms of speech accent across different levels of linguistic analysis using evidence from both behavioral and electrophysiological experiments. I further divide this research topic into three specific research questions distributed throughout the following chapters in order to address the challenges mentioned in section 1.2.1.2.

Linguistic challenges:
RQ1. How does accent affect early processing mechanisms during isolated single-word processing?
Cognitive challenges:
RQ2. How does accent affect the listening effort and memory of sentence processing?
Cognitive-linguistic integration challenges:
RQ3. How does accent affect processing mechanisms during extended discourse processing?

2.4.2. Overview of the experiments

RQ1, the focus of chapter 3, uses EEG ERP evidence to differentiate between early, bottom-up processing mechanisms in native, dialectal and foreign accents. The linguistic challenges of foreign and dialectal speech are addressed in RQ1, where I have explored early processing mechanisms of accent at the phonological and lexical-semantic level of single-word processing. This study aimed to disentangle the competing hypotheses of dialectal accent processing: the Different Processes Hypothesis and the Perceptual Distance Hypothesis. Electroencephalographic data was recorded from Spanish-Basque early bilinguals from the Basque country who listened to Spanish isolated words in Spanish in different accents (native, dialectal, foreign).

RQ2, the focus of chapter 4, shifts the focus of the investigation from single words to sentences. RQ2 aimed to address the cognitive challenges of accented speech processing, exploring the attentional and memory effects of dialectal and foreign accent as the linguistic level of processing becomes slightly more complex. Using behavioral evidence of a dual-task and recognition memory test, this study examines the effects of accent on listening effort, on memory and on their interaction. Specifically, I was interested in whether more attentional listening effort is required to process foreign-accented speech and importantly, whether any differences in memorization between information received in different accents was due to differences in listening effort. Participants followed a dual-task paradigm, listening to Spanish sentences in multiple accents (native, dialectal, foreign) in a blocked design while completing a secondary task (circle tracking on the screen). For each accent, listening effort measures were collected together with memory scores in an immediate and in a secondary session (4-6 days).

RQ3, the subject of chapter 5, explored the challenges related to the integration of cognitive and linguistic processing due to the interaction of bottom-up and top-down mechanisms that the longer time course of fluid discourse allows for. In this chapter, we return to the electrophysiological domain to examine the effects of accent on the top-down mechanisms of discourse processing. This study further addresses the competing hypotheses of dialectal accent processing from the frequency domain of EEG data, by examining in which frequency bands native, dialectal and foreign speech processing differ and where dialectal accent is placed with respect to foreign accent.

2.4.3. Significance and implications

Taken as a whole, this dissertation aims to contribute to the literature surrounding accented speech processing. With increased international mobility and the development of artificial intelligence, people are interacting with foreigners, and even synthetic voices, at an unprecedented rate. Critically, interlocutors in essential information-disseminating roles-such as our doctors, teachers and co-workers- are increasingly likely to come from a different background than us. People, all over the world, will now be increasingly attempting to remember information from speakers moving away from “standard” in-group members, such as speakers with foreign accents. The more we know how linguistic and cognitive processes (such as attention and memory) are affected by speakers’ characteristics, such as accent, the more we can prepare ourselves to understand and be understood in a global world.

3. Chapter III: An ERP investigation of accented isolated single word processing

3.1. Rationale

The first study that I will present in this investigation aims to take advantage of the precise temporal resolution of ERP analysis to explore online processing and to define the ms-by-ms time course of accented speech comprehension at the word level. Another advantage of this technique is the existing literature, which provides clear and well-established anchor points that we will use as a basis for our hypotheses. For instance, PMN and N400 ERP modulations during sentence comprehension are well-documented findings with consistent results across languages, so these well-known effects in native speech will be the baselines to understand how much sentence processing varies when speech is accented. Thus, the choice of the ERP technique makes this project innovative but also feasible and well-anchored to previous knowledge in the field. The information provided in this chapter has been published in the journal Neuropsychologia (Thomas et al., 2022).

3.1. Introduction

As outlined in the introductory chapters, it is well documented that unfamiliar accents can lead to difficulties in speech comprehension (Munro & Derwing, 1995a; Munro & Derwing, 1995b; Schmid & Yeni-Komshian, 1999; Anderson-Hsieh & Koehler, 1988; Major et al., 2002). It has furthermore been shown that non-native speech can lead to even greater difficulties (Floccia et al., 2006). As conversations with foreign and dialectal accented speakers become increasingly likely, it is important to understand the processing mechanisms involved in these non-standard types of speech comprehension.

Within the realm of accented speech research, much of the research has focused on non-native, or foreign, accented speech processing. However, studying accents in a native, non-native binary does not reflect the true level of complexity of cross-cultural interactions. Often, we may also speak with people who share our native language but are from a different country or region than us. These accent types, dialectal accents, would have native phoneme variations, unlike foreign accents, and thus it is important to understand the spectrum of accent processing and how processing mechanisms may differ between them. Two hypotheses have been proposed to compare processes underlying foreign and dialectal accent adaptation.

The Perceptual Distance Hypothesis relies upon the knowledge that accents can be placed on a perceptual scale based on their acoustic distance from native speech, foreign accents being the furthest from native and dialectal accents falling somewhere in between. According to this hypothesis, dialectal accent processing mechanisms would be attenuated versions of foreign-accented processing mechanisms (Clarke & Garrett, 2004). Behavioral evidence for this hypothesis was reported by (Adank et al., 2009), who found that processing speed decreased as accents were less familiar or foreign. (Floccia et al., 2006) similarly favored this hypothesis when they found that spoken words were recognized more slowly when produced in an unfamiliar accent, and even more slowly in a foreign accent, as compared to the participants' native accent.

However, other studies have found data inconsistent with this theory, including additional studies by Floccia (Floccia et al., 2009); Girard et al., 2008). This led to the proposal of the Different Processes Hypothesis, which posits that there are qualitative differences in the processing mechanisms recruited for dialectal and foreign speech normalization (Floccia et al., 2009). Studies such as Floccia et al. (2009) and Girard, Floccia & Goslin (2008), mentioned above, have supported this hypothesis by showing that French-speaking children (Girard et al., 2008) and English-speaking children (Floccia et al., 2009) are not able to reliably distinguish between dialectal and native accents while they are able to distinguish between foreign and native accents, even when accents strengths between the non-native accents are equated, suggesting differences in the linguistic characteristics of the respective speech signals and required normalization mechanisms.(Evans & Taylor, 2010) also supported the idea of differential processing for foreign and dialectal accents by showing easier adaptation to foreign than dialectal accents, despite slower and more error-prone responses with the foreign accents, in a comprehension-in-noise task.

Along with behavioral techniques for studying accented speech comprehension, electrophysiological methods have been increasingly employed due to their ability to provide valuable time-sensitive information. Within this technique, several event-related potentials (ERPs) have been found to be related to various processes during spoken word recognition.

One such component of interest is the P2 or P200 component, which is a positive-deflecting, anterior waveform peaking at around 200 milliseconds (ms) post-stimulus and often distributed around the centro-frontal area of the scalp. This component is associated with the extraction of acoustic features in speech (Reinke et al., 2003). The amplitude of the P2 increases with ease of extraction and is smaller in situations where acoustic extraction is more difficult. In the context of accented speech, a reduction in the amplitude of the P2 has been observed with foreign-accented speech reflecting an increased difficulty in extracting phonetic information (Romero-Rivas et al., 2015). This study did not include a dialectal accent contrast so it is difficult to say which hypothesis their findings support. Another study that included both a dialectal and foreign accent, aimed to investigate the effects of in-group membership and confidence on trust (Jiang et al., 2020). They found a reduction in the P2 only for the dialectal accented condition. However, they note in their discussion that while the native and foreign accents used (Canadian accent and Quebec French accent) were highly familiar to the participants due to the geographical region of the data collection, the dialectal accent (Australian) was much less familiar, which could have influenced the results (Jiang et al., 2020).

Another component of interest is the Phonological Mismatch Negativity (PMN), a negative-going waveform that generally peaks between 250 and 300 ms post-stimulus onset and is usually distributed along a centro-frontal line. The PMN is related to the normalization between acoustic input and phonological representations (Newman & Connolly, 2009; Goslin et al., 2012). This component is thought to reflect pre-lexical level processing (Connolly et al., 2001) (Newman et al., 2003). Phonological incongruencies, or a mismatch between perceived speech input and a lexical representation, are thought to elicit a larger PMN (Connolly & Phillips, 1994). Because of this finding, the PMN has been described as representing a ‘goodness-of-fit’ marker between phoneme representations and spoken input (Newman & Connolly, 2009). The interpretation of the PMN as a ‘goodness of fit’ marker, is in line with the Perceptual Distance Hypothesis because it supports the idea that the PMN is less pronounced as speech input gets close to native pronunciations (e.g., dialectal or highly proficient non-native speech). However, other recent evidence has argued that the PMN may be better characterized as an index of expectation, rather than an index of phonological mapping, due to their findings that accent experience modulated PMN effects possibly due to experienced participants’ adjusting expectations for non-native speech (Porretta et al., 2017). Other studies such as Goslin et al., 2012 and Sumner & Tilsen, 2011 have supported the Different Processes Hypothesis by providing evidence that dialectal and foreign accents recruit different normalization mechanisms. For example, Goslin et al., 2012 found that the amplitude of the negative-going PMN for dialectal accents was significantly greater than that of the native accent while the PMN for foreign accents was significantly smaller. This led them to conclude that different strategies were employed to process dialectal and foreign accents at early processing stages (200-350ms).

Finally, the N4 or N400, a negative-deflecting waveform that usually peaks around 400 ms post-stimulus with a widespread distribution that generally reaches maximum negativity at centroparietal sites, is also of interest when studying spoken word recognition. It has been associated with auditory semantic processing (Bentin et al., 1993; Kutas & Hillyard, 1980) and lexico-semantic integration (Kutas et al., 1987). There have been mixed results on the effects of accent on the N400. Hanulíková et al., 2012 observed a similar N400 effect during listening to both native and foreign-accented spoken sentences with semantic violations (cf. Gosselin et al., 2021). While Goslin et al., 2012 found a reduced N400 effect in correctly spoken sentences during foreign accent listening as opposed to dialectal or native listening. Goslin et al., 2012 thus argued that participants employed top-down resources to successfully perform word recognition when comprehension was degraded due to foreign accent (relative to dialectal and native). They argued that their results provided further evidence for the Different Processes Hypothesis due to the ability to successfully normalize dialectal accents during pre-lexical processing while foreign accents could not be normalized and thus continued to have an effect at lexical access and integration stages.

While there is strong evidence that foreign accent processing is different from native accent processing, it is still not clear how dialectal accents are processed nor the relationship between foreign, native and dialectal accent processing mechanisms. Previous studies have contributed to the plausibility of both previously introduced hypotheses, and importantly for our purposes, previous ERP experiments exploring the question have used sentences as material. More studies are needed to elucidate whether foreign accent is processed uniquely from all types of native speech (both native and dialectal accents) or whether dialectal accent is treated differently from native accent, despite both being native speech variations. Exploring accent processing in these three accent types through isolated word listening is useful to clarify accent processing mechanisms without the influence of sentence context, which adds complexity through potential prediction mechanisms and top-down effects. Such a low-level analysis at a small speech unit is convenient to provide a clear, simple picture to advance the debate on the Perceptual Distance and Different Processes Hypotheses. Thus, this will be the focus of the present study.

3.1.1. The Present Study

The present study aims to investigate the processing mechanisms of accented speech using electrophysiological methods. We hope to clarify the processing of foreign, dialectal and native accented speech through the lens of previously proposed hypotheses about accent processing. Specifically, we aim to reveal more about how we process dialectal accent and whether it is treated more similarly to a foreign accent or a native accent during isolated word recognition. ERPs were used because their rich temporal resolution was especially appropriate to study the time course of isolated-word processing mechanisms.

If results support the Perceptual Distance Hypothesis, dialectal accent processing should be more similar to foreign accent processing, although with attenuated effects. Thus, according to the Perceptual Distance Hypothesis, we should observe a gradient in the ERP responses in one or more ERP components that are crucial for speech recognition.

In contrast, a situation where the brain is mainly sensitive to a native/non-native distinction would be in line with the Different Processes Hypothesis, meaning that foreign-accented speech will be treated differently from both dialectal and native accent processing.

3.2. Methods

3.2.1 Participants

Thirty Spanish natives[4] participated in the study. Five participants were excluded from further analyses due to excessive artifacts in the electroencephalogram (EEG) recording or due to data collection error (missing behavioral data of one participant). The final sample of participants

[4] Due to their geographical location in the Basque Country, all participants were also fluent in Basque as early L2

consisted of 25 females[5] (mean age: 23 years, SD: 3.54, age range: 19-31 years, Spanish age of acquisition: 0). A post hoc power analysis of the final sample (n=25), using the software G* Power (Erdfelder et al., 1996) revealed the power to detect a medium-sized effect (f=0.25; cf (Cohen, 1977) with alpha at 0.05 was 0.76. All participants lived in the Basque country and considered the Basque-Spanish accent as their native accent. All participants were right-handed and had normal or corrected-to-normal hearing and vision. No participant reported a history of neurological disorders. All participants signed an informed consent form before taking part in the study that was approved by the Basque Center on Cognition, Brain and Language ethics committee. They received monetary compensation for their participation.

3.2.2. Materials and Methods

Forty Spanish words were extracted from dialogues recorded by 6 female speakers in their thirties with differing accents and presented three times (tot: 120 words per accent, see Appendix A). The words were recorded in a Basque Spanish native accent, a Cuban Spanish accent and an Italian Spanish accent by two females for each accent (i.e., hereafter native, dialectal and foreign accent, respectively). The native speakers were born and lived in Spain (Basque country). The dialectal and foreign speakers were chosen for their strong accents and high level of Spanish (in the case of the Italian speakers). Overall, the recordings did not differ in duration (ms) across accents [foreign: 519.7, SD:128.7; native: 505.7, SD: 137.2; dialectal: 519.9, SD:131.6; one way ANOVA: ($F(2,117)=0.15$, $p=0.86$)]. Accent strength ratings were collected from a separate normative study consisting of eleven participants (average age =23, SD=10) who completed a short online survey where they listened to clips of each accent and rated the accent strength from 1 (mild accent) to 5 (strong accent), showed a clear effect of accent (one way ANOVA: $F(2,10)=6.01$, $p=0.009$). Follow-up analyses of the clip ratings corrected with the Fisher's Least Significant Difference (LSD) showed that the dialectal accent was marginally significantly different than the native accent ($t(10)=2.13$, $p=.06$), the foreign accent was significantly different from the native ($t(10)=3.16$, $p=.00$) and the dialectal and

[5] Only females were recruited for this study in order to avoid cross-gender listening effects (Banaji & Hardin, 1996; Cacciari & Padovani, 2007)

foreign accents were not rated significantly different from each other in terms of strength ($t(10)=1.15$, $p=.28$).

3.2.3. Procedure

Participants were seated in a sound-attenuated room in front of a computer screen and were asked to listen to words and occasionally repeat the last word they heard when instructed to by writing on the screen (to maintain listeners' attention). Each trial began with the symbol *.* at the center of the screen, which was followed by a 500-msec fixation cross. Then a word was presented through speakers while a fixation cross was displayed on the screen. Thirty-three percent of the words were followed by the Spanish word for repeat ('repite') displayed on the screen to indicate that the participant should repeat the word that they heard. To minimize artifacts in the EEG recording during the presentation of the auditory stimuli, the participants were asked to blink only when the symbol *.* was presented on the screen. The experimental session lasted about an hour and was divided into 9 blocks. We adopted a blocked design where each accent type was presented in 3 consecutive blocks. The accent presentation order was counterbalanced across participants. Within each block, 40 words of the same accent type were presented in a randomized order (20 words for each speaker).

3.2.4. EEG Data Recording and ERP Analyses

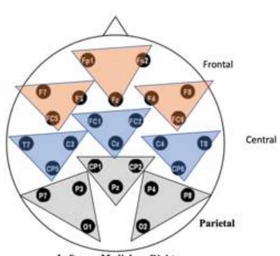


Fig.6. Schematic of electrode montage with topographic organization labeled.

The EEG signal was recorded from 27 channels placed in an elastic cap: Fp1, Fp2, F7, F8, F3, F4, FC5, FC6, FC1, FC2, T7, T8, C3, C4, CP5, CP6, CP1, CP2, P3, P4, P7, P8, O1, O2, Fz, Cz, Pz (see Fig. 6). Two external electrodes were placed on the mastoids, two were on the ocular canthi, one above and one below the right eye. All sites were referenced online to the left mastoid. Data were recorded and amplified at a sampling rate of 250 Hz. Impedance was kept below 10 KΩ for the external channels and below 5 KΩ for the electrodes on the scalp. EEG data were re-referenced offline to the average activity of the left and right mastoid. A low-pass filter of 30 Hz and a high-pass filter of 0.01 Hz were applied. Vertical and horizontal eye movements were corrected following the Independent Components Analysis (ICA). The fast ICA-restricted method was used (1470.3 s). The EEG of each participant was decomposed into independent components and components that explained the highest percentage of the variance in the Veog and Heog channels were identified. Data inspection was then performed on these components to ensure they were accurately accounting for variability related to ocular artifacts. Residual artifacts exceeding ± 100 uV were rejected. On average, 13% of trials were excluded. The number of rejections did not differ across conditions (one-way ANOVA: $F(2,72)=0.20$, $p=0.82$). For each target word, an epoch of 1200 msec was obtained including a 200 msec pre-stimulus baseline. Average ERPs time-locked to the word onset were computed and baseline corrected for each condition. Statistical analyses were carried out in the following time windows: 150-250, 250-400, and 400-600 msec. The temporal boundaries of each time window were defined based on visual inspection and were also similar to those used in previous ERP studies on auditory word comprehension (Hanulíková et al., 2012; (Rossi et al., 2006; Schmidt-Kassow & Kotz, 2008); Van Berkum et al., 2008).

Repeated measures analysis of variance (ANOVA) using conservative degrees of freedom (Greenhouse & Geisser, 1959) was run on the average waveform amplitudes within the time windows of interest. Each ANOVA consisted of accent as a three-level factor (native, dialectal, foreign) and included a topological organization of longitude (frontal, central, parietal) and latitude (left, medial, right). All electrodes were considered in the ANOVA in order to enhance statistical power and directly test the distribution of the effects without theoretical constraints. Each longitude x latitude combination was linearly derived from a combination of three sites: left frontal (F7, F3, FC5), medial frontal (FP1, FP2, Fz), right frontal (F4, F8, FC6), left central (T7, C3, CP5), medial central (FC1, FC2, CZ), right central (C4, T8, CP6), left parietal (P7, P3, O1), medial parietal (CP1, CP2, Pz), right parietal (P4, P8, O2). Finally, t-test post hoc comparisons were conducted and corrected for multiple comparisons using the Benjamini-Hochberg procedure. Additional analyses in the time-frequency domain were also run and confirmed the ERP results, see Appendix B for more details.

3.3. Results

3.3.1 Behavioral Results

Participants showed high accuracy during the online word repetition task meant to test attention and intelligibility (mean overall accuracy: 93.8%, SD:6.8). The word repetition scores did not differ between accents [foreign: 95.3%, SD:10.3; native: 93.6%, SD: 7.6; dialectal: 92.6%, SD:10.0; one way ANOVA ($F(2,72)=0.53$, $p=.59$)] showing that the participants could perceive and repeat the words equally well in every accented condition.

3.3.2. ERP Results

Average ERPs time-locked to 200 ms pre-word onset in native, dialectal, and foreign accent conditions can be seen in Fig. 7. Topographic distribution for each accented comparison that reached significance was calculated in the 150-250 ms and 250-400 ms time windows and can be seen in Fig. 8 and 10. Plots showing the individual data points and topographic distribution of data for each time-window of interest can be seen in Fig. 9, 11, and 12.

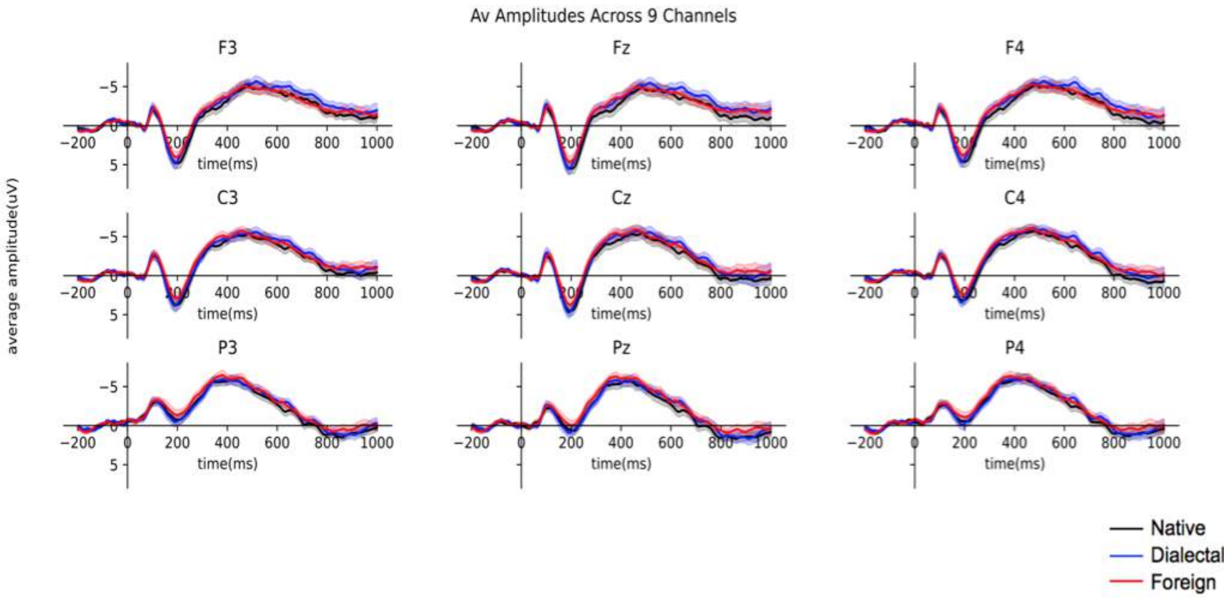


Fig. 7. Grand average ERP amplitudes from 200ms pre-word onset until 1000ms after, in different accents. Negativity is plotted up. One standard error is shown.

2.3.2.1. P2 [150-250 ms]

A main effect of accent was found ($F(2,48) = 5.19$, $p < .05$, $\eta^2=0.34$), with a reduced P2 amplitude only for foreign accent (foreign vs native: $t(24)=2.92$, $p<.01$, foreign vs dialectal: $t(24)=2.88$, $p<.01$, dialectal vs native: $t(24)=0.37$, $p=0.71$). The interaction between accent and longitude was also significant ($F(4,96) =4.03$, $p<.05$, $\eta^2=0.33$) suggesting that the difference between foreign and native accent was anteriorly distributed ($t(24)=3.36$, $p<.05$), while the difference between foreign and dialectal accent was posteriorly distributed ($t(24)=3.14$, $p<.05$). All the other contrasts did not reach significance (ps>0.05).

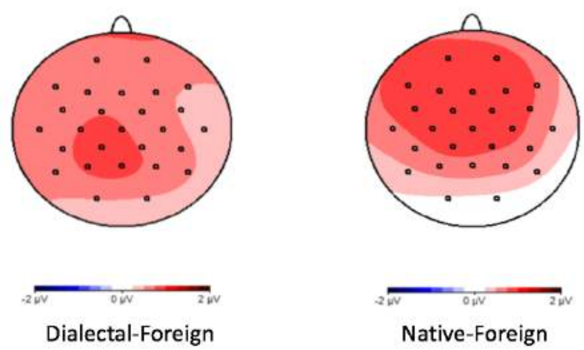


Fig. 8. Topographic distribution of voltage difference between conditions with significant differences between 150-250 ms.

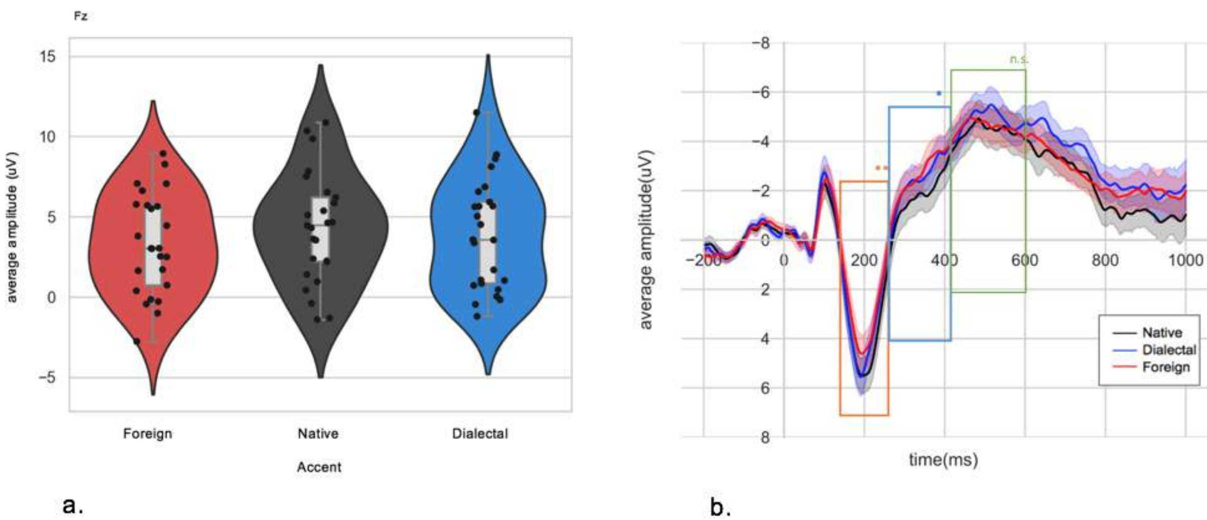


Fig. 9.
a. Violin plots of the average amplitude of each condition and individual scatterplots with significant differences between 150-250 ms, inner box plot is shown with median, third quartile and first quartile.
b. Grand average ERP amplitude from 200 milliseconds prior to word onset (-200) till 1000ms after, in different accents. Negativity is plotted up. The three time-windows of analyses are highlighted (orange for the P2, blue for the PMN, green for the N400). One standard error is shown.

3.3.2.2. PMN [250-400 ms]

A main effect of accent was found ($F(2,48) = 3.43$, $p < .05$, $\eta^2=0.30$), with the highest negativity in foreign accent, followed by dialectal and finally native accent (foreign vs native: $t(24)= -3.10$, $p<.01$, foreign vs dialectal: $t(24)=-2.67$, $p<.05$). There was no significant difference between the native and dialectal amplitudes (dialectal vs native: $t(24)=0.59$, $p=0.56$). The interaction between accent and longitude ($F(4,96) =3.52$, $p<.05$, $\eta^2=0.28$) and accent, longitude and latitude ($F(8,192)= 2.02$, $p<.05$, $\eta^2=0.60$) was also significant suggesting that the effect of accent was mainly distributed anteriorly for the foreign-native difference ($t(24)= -3.59$, $p<.01$) and posteriorly for the foreign-dialectal difference ($t(24)= -3.03$, $p<.01$). The posterior effect in the foreign-dialectal contrast was somewhat left-lateralized ($t(24)= -3.22$, $p<.01$).

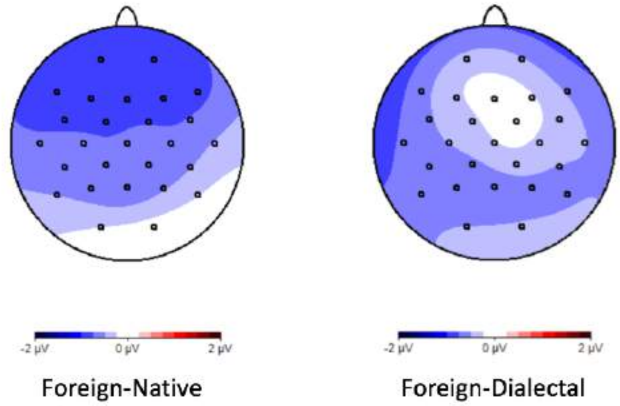


Fig. 10. Topographic distribution of voltage difference between conditions with significant differences between 250-400 ms.

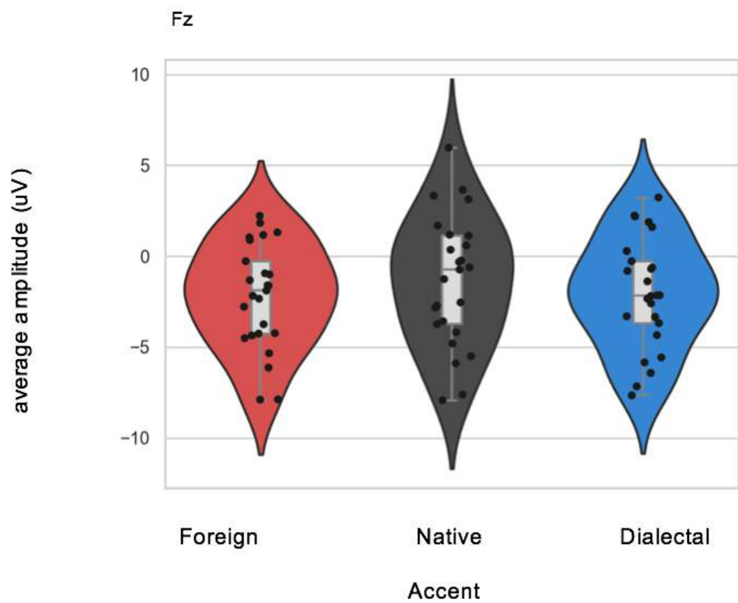


Fig. 11. Violin plots of the average amplitude of each condition and individual scatterplots with significant differences between 250-400 ms, inner box plot is shown with median, third quartile and first quartile.

3.3.2.3. N400 [400-600 ms]

No significant effect of accent was observed ($F(2,48) = 0.29$, $p = .72$, $\eta^2=0.02$). Additionally, no significant interactions with accent were found (accent x longitude: $F(4,96)=0.60$, $p = .41$, $\eta^2=0.12$; accent x latitude: $F(4,96)=0.79$, $p = .50$, $\eta^2=.18$; accent x longitude x latitude: $F(8, 192)=0.62$, $p=.12$, $\eta^2=0.29$).

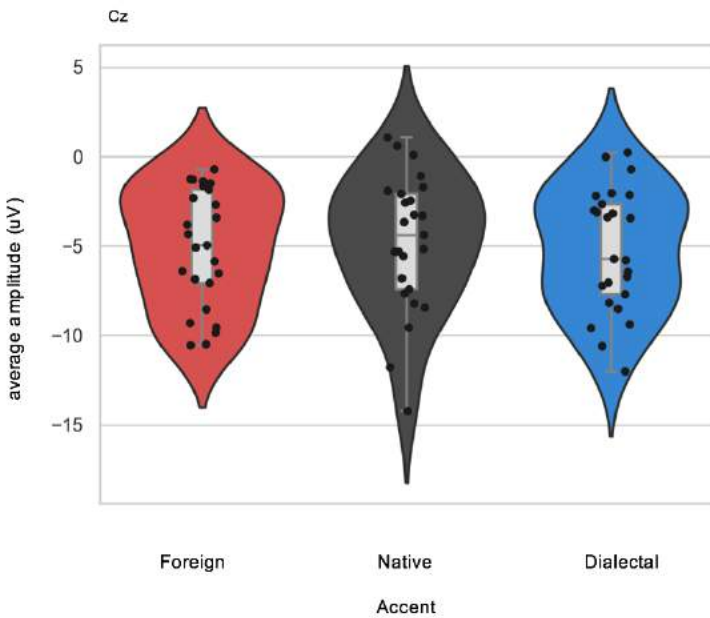


Fig. 12. Violin plots of the average amplitude of each condition and individual scatterplots between 400-600 ms, inner box plot is shown with median, third quartile and first quartile.

3.4. Discussion

This study aimed to add to the literature on the mechanisms of accented listening comprehension by attempting to advance the debate on two popular spoken language processing theories: the Different Processes Hypothesis and the Perceptual Distance Hypothesis. We were especially interested in contributing to the literature on disentangling dialectal and foreign accents.

Previous literature has shown processing differences between native and non-native accented speech listening (Lane, 1963; Lev-Ari & Keysar, 2012, 2017; Munro & Derwing, 1995a, b; van Wijngaarden, 2001) and to a lesser extent native and dialectal variations (Adank & Janse, 2009; Floccia et al., 2006; Floccia et al., 2009; Goslin et al., 2012; Girard et al., 2008). These differences have led researchers to debate whether accent processing exists on a perceptual scale based on acoustic nearness to native speech (Perceptual Distance Hypothesis) or in distinct categories (Different Processes Hypothesis).

We examined this issue by looking at the effect of accent on the modulation of the P2, PMN and N400 ERP components while listening to isolated words in three accented conditions (foreign, dialectal and native).

Our results align with the Different Processes Hypothesis supporting the idea that foreign-accented speech, at the isolated word level, may present a unique processing challenge that is not present in dialectal speech processing. It further appears that the processing costs of accented speech are strongest early on during the time course of isolated word listening and seem to get weaker over time, suggested by the lack of an effect of accent after 400 ms.

We found that P2 is modulated by accent, with a reduced P2 for foreign accent. This result is in line with previous findings that early stages of speech comprehension are compromised when processing foreign-accented speech due to the phonological properties of foreign-accented speech often departing from those of the native listener (Lane, 1963; Munro & Derwing, 1995b; van Wijngaarden, 2001). Our results in the P2 were similarly congruent with the conclusions of Romero-Rivas et al., 2015, suggesting that the extraction of acoustic features from foreign-accented speech is more difficult than from native. Our results furthered this conclusion by also determining that the extraction of acoustic features from foreign-accented speech was also more difficult than from dialectal, while the extraction of both dialectal and native acoustic features was similarly easy. Therefore, ERP results at the P2 support the Different Processes Hypothesis due to the reduction of the P2 only observable in the foreign-accented condition. This leads us to believe that extracting non-native speech sounds presents a unique challenge that is not triggered by native but non-local speech sounds (i.e. dialectal). This processing effort was not seen for our dialectal condition despite the fact that the Cuban dialectal accent used was considered strongly accented by our normative study participants and phonologically deviated significantly from our native accent of Spanish. This may mean that the 'coherent deviations' (see Wells, 1982; Goslin et al., 2012) that are present in dialectically accented speech are easily adapted to by listeners and thus do not interfere with the extraction of spectral information and other acoustic features, while foreign deviations make this extraction more difficult.

We also found that the PMN is modulated by accent, with the largest mismatch negativity elicited from foreign accent, reflecting increased resources required by normalization processes. This finding is partially in line with previous conclusions drawn about the PMN that support the Different Processes Hypothesis, such as Goslin et al., 2012, who compared the ERPs associated with the perception of fully formed sentences spoken in different accents. They found that the amplitude of the PMN elicited by the foreign accent and that by the dialectal accent were significantly different from each other. However, our findings diverged from theirs in the directions of amplitudes found and a lack of significant difference in the native-dialectal comparison. Whereas Goslin et al., 2012 found that the amplitude of the PMN for dialectal accent was significantly greater than that of the native and the foreign was significantly smaller than that of the native, we found a gradient in amplitudes moving from native to foreign, with foreign accent eliciting the greatest negativity. Thus, our findings are more in line with the idea of the PMN as a ‘goodness of fit’ marker. Although numerically we did observe a gradient in the PMN averages of our three accented conditions, the difference between the PMN elicited by native and dialectal accents was not significant thus this study did not provide evidence supporting the Perceptual Distance Hypothesis. However, because the trend of the amplitudes in the PMN was consistent with the Perceptual Distance Hypothesis, we conducted an exploratory Bayesian analysis in order to elucidate a clearer picture of pre-lexical accent processing mechanisms and verify the reliability of the null results. This analysis revealed that there was a strong effect between the foreign and other two accents (BF= 1.6e+9; BF =1.3+5) while there was moderate evidence for the null hypothesis (i.e., lack of difference between native and dialectal accent) in the PMN (BF = 0.42), adding further credibility to support of the Different Processes Hypothesis.

Finally, by the 400-600 ms epoch, we no longer observed any significant effect of accent. This is, again, partially in line with Goslin et al., 2012 who in the 350-600 ms epoch no longer observed a difference between native and dialectal accent conditions but diverged from their finding of a significant difference in the foreign accent condition. While their study suggested that “non-native phonological variations could not be fully normalized in pre-lexical processing levels, and so had a continued effect at lexical access and integration stages” (pp.102), our results suggest that in isolated word listening, participants are able to normalize foreign accent at the pre-lexical processing level and thus the accent no longer affects the lexical access stage. These differences are most likely due to using isolated words instead of sentences and may indicate flexibility of this processing mechanism.

The results of this study have suggested that non-native accent affects early stages of single-word processing. They have further suggested that these effects are uniquely seen for non-native accents rather than also native variations. While our results suggest that at the isolated word level the brain is operating on a native/non-native binary distinction, the question remains about how general this apparently binary mechanism is. Is it truly a generalized cognitive mechanism or perhaps the flexibility of processing is influenced by external factors? One factor that might be relevant is the acoustic distance among different accents. The acoustic features of dialectal accents might be less salient than those of foreign accents. Additional research can clarify the role of acoustic properties in accent speech perception of isolated words. Future studies can try to generalize these results to different types of participant profiles, based on gender, language, exposure to accents, or other factors of interest. In addition, would we observe a dichotomic treatment of native/non-native accents in situations involving much more top-down processing, such as continuous speech? Chapter 5 will explore the role of top-down processes and suprasegmental information in continuous speech to clarify the role of this mechanism. The impact of additional factors, such as acoustic distance and top-down mechanisms, could reconcile differences seen in our study and other studies conducted using sentences that contain additional suprasegmental (e.g. prosody stress, tone) information.

3.5. Conclusion

Our results support the idea that, even during mostly passive isolated-word listening, indexical features, such as accent, affect our processing at early stages. Furthermore, our findings support that foreign accents are processed differently from dialectal or native accents, requiring more effort to extract acoustic characteristics and to normalize them into expected phonological representations. Despite this, it appears that we are able to successfully normalize non-native speech and thus by later processing stages accent no longer affects our processing or ability to integrate speech sounds into lexical information. Because it remains unclear if we would observe this dichotomic treatment of native/non-native accents in situations involving much more top-down processing, the next chapters will focus on longer units of speech.

4. Chapter IV: A dual-task investigation of the effects of accented sentence processing on attention and memory

4.1. Rationale

In the last chapter we examined online processing differences at the small linguistic unit of a single word spoken in isolation. Now, we shift to behavioral methods in order to examine the cognitive load of more natural speech. Thus while the last chapter focused on isolated words, this chapter increases the level of linguistic analysis to that of the sentence. The chapter examines the effects of accent on listening effort, on memory and on their interaction. Millions of people interact on a daily basis with non-native speakers of their language. While we are increasingly likely to have conversations with foreign speakers, it has been shown that attending to foreign-accented speech presents us with unique difficulties (Lev-Ari & Peperkamp, 2014). Furthermore, although research has indicated that accent can impact attention and memory representations (Lev-Ari, 2017; Atkinson et al., 2005), the specific causal relationship between these factors remains relatively unexplored. Moreover, there is limited knowledge regarding the long-term effects of foreign accents on an individual's ability to retain information provided by a non-native speaker over an extended period. The objective of the study outlined in this chapter was to investigate the influence of various accents on cognitive processing and the ability to remember information over an extended period.

4.2. Introduction

In our current globalized and technologically connected society, we are more likely than ever to receive information from a wide range of speakers. With rising international mobility, people are interacting with foreigners at an unprecedented rate. Critically, interlocutors in essential information-disseminating roles-such as our doctors, teachers and co-workers- are increasingly likely to come from a different background than us. People, all over the world, will be increasingly attempting to remember information from speakers moving away from "standard" in-group members, such as speakers with foreign[6] or dialectal[7] accents. Do we

[6] linguistic output which possesses acoustic features that deviate from the native pronunciation of a given language
[7] Linguistic output which possesses native phoneme variations

struggle to attend to and memorize the details of a doctor's instructions when the doctor has an unfamiliar accent? Do we forget a lesson faster if taught by someone from abroad? Because of these questions, it is beneficial to know how cognitive processes such as attention and memory are affected by a speaker's accent. These concerns are especially relevant in bilingual environments, where people are especially immersed in speaker-specific speech variations. This project investigates the impact of a speaker's accent on the attentional load and information retention of speech processing by testing listening effort, speech retention processing and information retention over time.

Previous research has found that speaker-related characteristics can have an impact on a range of offline and online cognitive processes (Baus et al., 2017; McAleer et al., 2014; Stevenage et al., 2012; Sporer, 2001). For instance, interlocutor identity may place different attentional demands on the listener due to differences in intelligibility (Oppenheimer, 2008), social group (Sporer, 2001) etc. Particularly salient examples of this are speaker-related features (e.g., accent type).

Previous studies have demonstrated that foreign-accented speech may be more difficult to attend to than native-accented speech (Lev-Ari & Peperkamp, 2014). Foreign accented speech listening has been linked to increased processing time, resulting in longer response latencies when evaluating the truth value of statements even when comprehensibility is high (Munro & Derwing, 1995b) and slower detection of the mispronunciation of words (Schmid & Yeni-Komshian, 1999). Processing foreign-accented speech requires the listener to adapt to an extra range of variability, suggesting that there may be an increase in the amount of our attentional and cognitive resources that are needed to successfully interpret the speech signal of a foreign-accented speaker. Indeed, recent studies have found that the attentional demands of foreign-accented listening appear to be higher than native-accented listening (Yi et al., 2014; Brown et al., 2020b).

In addition to fluctuations in attentional demands, it seems that speaker characteristics can lead to differences in the depth of listeners' memory encoding as well. One study showed that performance on comprehension questions decreases when listening to either a computer-generated voice or a foreign-accented speaker, as compared to a native-accented speaker (Atkinson et al., 2005). Additional research in the domain of accent processing has found that listeners have less detailed representations of the speech of non-native speakers

(Lev-Ari & Keysar, 2012; Lev-Ari, 2017), while dialectally accented speech does not seem to impact the memory or credibility of the message (Frances et al., 2018).

Despite these findings, less research has been done on investigating the memory effects of speaker-type characteristics. Further exploration within this field is necessary to deepen our understanding of how accent affects our memorization and retention of information. For instance, familiarity and recollection paradigms[8] have been frequently employed in recognition memory research. This is because it offers the possibility to differentiate between the theorized separate memory processes of the experience of recognizing something as familiar but not being able to recollect details and being able to retrieve details associated with the previously experienced event (for a review see Yonelinas, 2002). While these paradigms have been employed in studies of novel word list encoding and retention, less is known about recognition memory categorization in extended listening tasks. In addition, despite knowing that accent can lead to differences in both attention and memory representations (Lev-Ari, 2017; Atkinson et al., 2005)), little is known about the possible causal relationship between them or the extended consequences of a foreign accent on one's ability to remember information. Thus, this study aims to examine the effect of different accents on cognitive demands and subsequent information retention.

Based on the gap in this literature, three different models, within the domain of the allocation of cognitive resources and subsequent effects on memory retention, have been proposed: the Cognitive Load Theory, the Cognitive Effort Theory and the Indexical Free Theory, as discussed in the general introduction. The first of these theories is the Cognitive Load Theory (see Paas, Tuovinen, et al., 2003 for an overview). Cognitive load refers to the amount of cognitive demands that are placed on the working memory system in order to perform a given cognitive task (Pichora-Fuller et al., 2016). Similarly, cognitive effort is a term that refers to the amount of cognitive resources that one allocates to perform a task (de Jong, 2010). Cognitive Load Theory hypothesizes that an attentionally or cognitively demanding situation would lead to a drop in memorization performance (Paas et al., 2016; Paas, Renkl, et al., 2003). This is because cognitive resources are inherently limited. As the required cognitive effort increases, the cognitive load is assumed to be greater as well. The increased cognitive load is thought to have

[8] recognition memory test asking whether something was present in previously presented material vs asking for specific details about the previously encountered item

negative consequences on one's total cognitive capacities thus making memory encoding and retrieval more difficult.

Conversely, the Cognitive Effort Theory hypothesizes that an attentionally demanding situation would lead to better memorization performance (for a summary of the theory, see Mayer et al., 2005; Tyler et al., 1979). According to this theory, an increase in the cognitive load of a cognitive task would lead to a reallocation or re-prioritization of cognitive resources due to the increased cognitive effort. This reallocation of resources would enable increased attention to a task, decreasing encoding and retrieval difficulty.

Finally, while the two previous theories are generalizable to many attentionally demanding cognitive processes, the third theory is specific to the scope of this study. The Indexical-Free Theory hypothesizes that accent is an attentionally demanding situation that should not affect memory performance (Romero-Rivas et al., 2015). This is because accent is considered an indexical property of speech, aspects such as age, gender, social status, native-accent status, or familiarity of the interlocutor. This theory suggests that indexical properties are removed before memory encoding or retrieval and therefore have no effect on these processes.

To investigate the aforementioned theoretical models, it is useful to look at the cognitive load across speaker accents to determine if possible performance differences are due to increased cognitive resource demands. One method for examining the cognitive load of a task is the dual-task paradigm. A dual-task paradigm is an experimental paradigm that requires a participant to complete two tasks of unequal importance simultaneously. The dual-task paradigm has proven to be a useful tool for quantifying listening effort, by providing a normalized measure of the cognitive load of simultaneously performing two tasks. Dual-task paradigms have been used to measure listening effort because a person's cognitive resources are limited in capacity and speed. Thus, once the maximum resources available have been reached, a decrease in performance on the non-prioritized task can be seen, reflecting increased listening effort (Gagné et al., 2017; Tun et al., 1991; Kemper et al., 2009). Although there are several methods of measuring listening effort, the dual-task paradigm is especially applicable to the present project due to the ecological validity of performing two tasks simultaneously, as we often must concentrate on multiple tasks when attending to speech, such as when driving, taking notes, etc (Gagné et al., 2017).

4.2.1. The Present Study

The present study used a dual-task paradigm where a primary (i.e., listening and repeating spoken sentences in a native, dialectal and foreign accent) and a secondary task (i.e., digital circle tracking) were provided to the participants. To examine the effect of the listening effort of different accents on information load and retention, participants were given a memory task, around 15 minutes after listening (session 1) and again 4 to 7 days later (session 2).

For the dual-task performance, based on previous evidence we predict that foreign accent should show a greater listening effort as compared to native accent, meaning that there would be an increased demand on attentional resources in the primary listening task and a consequently decreased performance on the secondary task of circle-tracking. For the memory task, different predictions can be made based on the higher attentional demands of the foreign accent. The Cognitive Load Theory predicts poorer memory retention performance for foreign compared to native accents, while the Cognitive Effort Theory predicts better memory retention performance for foreign compared to native accents. The Indexical-Free Theory, in contrast, predicts no difference in the memory retention performance for foreign compared to native accents.

As a secondary aim, the study introduces an additional manipulation to investigate the effect of dialectal accent on attention and information retention, teasing apart two competing theories on dialectal accent processing. While differences in attention have been shown for foreign-accented speech, the difficulty of attending to dialectal speech is less clear. These accent types would have native phoneme variations, unlike foreign accents, and thus it is important to understand how they may require a different amount of cognitive resources. Two hypotheses have been proposed to compare processes underlying foreign and dialectal accent adaptation. The Different Processes Hypothesis (Floccia et al., 2009) postulates that dialectal accent is part of the “native accent category” and therefore will show no increase in listening effort as compared to native speech. In addition, no effects on information retention should be observed. Conversely, the Perceptual Distance Hypothesis (Clarke & Garrett, 2004) proposes that there is a continuum of processing needs from native to dialectal to foreign accent, in this case, there will be an increase in the listening effort needed to attend to dialectal speech that is significantly different from both the native and foreign-accented speech conditions and an intermediate effect on information retention.

4.3. Methods

4.3.1 Participants

Forty-five Spanish natives[9] from the Basque Country, Spain (30 Female, mean age=26.24, SD=5.26, range=19-38) gave their informed consent and participated in the online experiment. Participants reported the Gipuzkoan accent as their native accent, matching the accent of the native condition recordings. All participants also reported normal hearing and normal or corrected to normal vision. The experiment was a two-session experiment. In the first session, participants completed the attentional measure, a short working memory task (N-back), and a subsequent memory exam. In the second session, participants completed a counterbalanced version of the memory exam that they previously had completed and a language background survey. Participants received monetary compensation at the end of both sessions. Participants were emailed the experiment links before each session. Four days after completing the first session, they were sent the second link and a reminder that they had until 7 days after completing the first session to complete the second session (average time between sessions:4.97 days, SD: 1.11). Overall, 9 participants were excluded due to failure to complete the second session (n=3), technical issue (n=4), or for not following instructions (n=2), making the final sample 36 participants (22 Female, mean age=25.89, SD=4.48, range=19-37).

4.3.2 Normative study on audio stimuli for accent strength ratings

17 participants (7 Female, mean age= 36, SD=13, range=23-61) who did not participate in the main study completed a short online survey. All participants listed their native language as Spanish and 3 listed Basque as a native language as well. No participants listed Basque as their dominant language. All participants were from Gipuzkoa. they listened to a sample of 18 audios, 2 sentences from each of the three speakers in each accent, therefore 6 sentences of each accent

[9] Due to their geographical location in the Basque Country, participants were also fluent in Basque as early L2

condition. Participants rated the accent strength from 1 (mild accent) to 5 (strong accent). Participants were then asked to try to identify where the accent was from. First, they had to choose between whether the accent heard was from Spain/Basque country, a different country whose native language is Spanish, or a different country whose native language is not Spanish. If they answered either of the two latter options, a third question appeared to ask them to type what country they believed the speaker was from. The mean accent strength ratings were approximately 3.9 for foreign (SD:0.61), 3.5 for dialectal (SD:0.094) and 2.0 for native accent conditions (SD:0.28).A one-way ANOVA with accent as a three-level factor determined that there was a main effect of accent on perceived accent strength ($F(2,48)=85.88$, $p=0.000$). Follow-up t-tests showed no significant difference between the perceived accent strength averages of foreign and dialectal recordings ($T(5)=2.13$, $p=0.08$) and a significant difference between the averages of the accent strength ratings for foreign and native($T(5)=7.35$, $p=0.000$) and also dialectal and native($T(5)=11.29$, $p=0.000$). Participants were overall very accurate at identifying whether the accents were from native Spanish speakers or not (all conditions around 95 to 98 percent accurate) but less accurate at identifying the country. For foreign-accented speakers, they were 58.5% accurate at identifying the correct country on average, for dialectal they were 72.5% accurate and for native they were 97% accurate.

4.3.3. Procedure and Methods

4.3.3.1. Session 1: DPRT task blocks

In the attentional session, while participants were listening to sentences they were also performing a digital pursuit rotor tracking task (DPRT), as in previous listening effort studies (Kemper et al. 2009, Desjardins and Doherty 2014, Peng and Wang 2019, Oosthuizen et al. 2020, Wu et al. 2014). The DPRT was designed to replicate as closely as possible the program used by Kemper et al., 2009, using javascript through the Pavlovia platform. The pursuit rotor featured a white circle on a black background that moved along an invisible elliptical track across the computer screen. When the cursor of the mouse was kept inside the circle, the color of the circle changed to green. Participants were instructed to use a mouse to track the circle. The size of the circle was determined by the resolution of the screen.

At the beginning of each DPRT block, the participant saw the white circle. Following an initial delay of 5 seconds in order to give the participant time to position the cursor for the first trial, the target began to move along the track. The participant attempted to keep the circle green by keeping the mouse inside the circle (see Fig. 13).
The software recorded for each frame the distance between the mouse pointer and the circle, and whether the mouse pointer was inside of the target circle. At the end of each DPRT block, the Time on Target (TOT) was determined, i.e. the percentage of time the mouse pointer was inside the circle for each condition. When the cursor of the participant was too far from the circle a warning message appeared in the center of the screen in red reminding the participant to follow the circle.

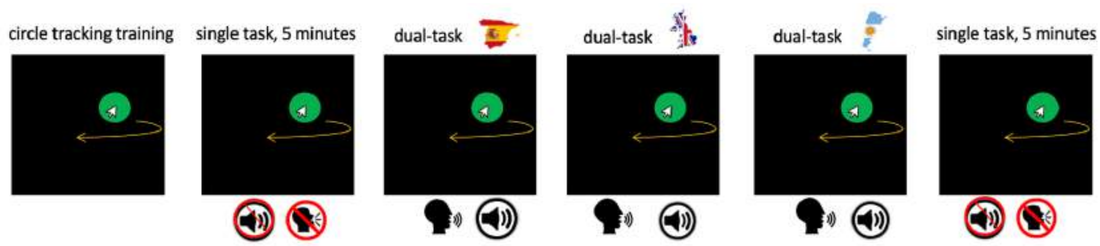


Fig. 13. Schematic of the attentional task, each dual-task block consisted of 25 sentences per accent.

4.3.3.2. Training

Participants began with a training block to set the speed of the circle to ensure they would be able to reach an asymptotic level of performance on the DPRT, thus avoiding floor or ceiling effects. This adaptive procedure was done as described in Kemper et al., 2009: participants had to track the circle for 30 seconds at a fixed speed. The initial revolutions per minute (RPM) for this phase was 3.4 RPM. After 30 seconds, the speed of the circle in RPM was modified according to the Time On Target (TOT). If the participant was more than 80% of the time with the mouse pointer inside the circle, the RPM was increased by 10% for the next trial. Otherwise, it decreased by 5%. This adaptive procedure finished when the participant crossed an RPM value three times going up and down. The final RPM was computed as the average between the final RPM, the last RPM that was higher than 80% TOT, and the last RPM that was lower than 80% TOT. The average RPM of the tracking speed of participants was 3.30 (SD: 1.39, Range: 1.04-6.50)

4.3.3.3. Single Task-Baseline One

After the asymptotic tracking speed was established for each participant, participants completed the initial baseline task, consisting of 5 minutes of tracking the circle without their attention divided on any other task. This same task was repeated after the 3 test blocks, to account for potential fatigue or learning effects.

4.3.3.4a. Dual Task

For the primary listening task participants listened to and repeated a total of 75 sentences, divided into three blocks by each accent type. Accent conditions were blocked so that the participant listened to and repeated 25 sentences of each accent before starting the next accent condition. Participants were told that they would now be listening to sentences and repeating them while continuing the tracking task of before. The program set the speed of the circle based on the asymptomatic tracking speed that was previously calculated for each participant in the training block. Though the tracking was continuous, trials were separated by beeps so that the participant was prepared for the next sentence to start and knew when to stop speaking. The screen prompted them to repeat after each sentence by displaying 'repite' (repeat) on the center of the screen. Five practice trials were included before the first test block to ensure that the participant understood the task. The participant was told that Google speech recognition would be annotating speech to text so they needed to speak as clearly as possible. During the practice trials, a native accented speaker was used to speak sentences that were not included in the actual experiment. The Google speech recognition software displayed the speech-to-text on the screen, allowing participants to adapt their speech in case of significant inaccuracies or if it was not understood properly. They were also told to stop the experiment and contact the experimenter if they did not see the text on the screen. They were informed that in the test blocks, they would no longer see the text on the screen. This was to avoid distraction and let them perform the tracking task without interference. In the case of the dual task blocks, TOT was calculated separately both for when the participant was listening to the stimuli and repeating the stimuli while completing the secondary tracking task.

4.3.3.4b. Recordings-Auditory Stimuli of the Dual Task

For the attentional session, three accented versions of Spanish sentences were recorded in a soundproof booth (sample frequency of 44.1 KHz). They were recorded in a Basque Spanish native accent, an Argentinian Spanish dialectal accent and a British (English) Spanish foreign accent. The recordings consisted of 75 sentences produced by three male speakers in each accent group (25 sentences each) in each of the three accented conditions (average duration=2,546 ms; SD=537 ms). In the case of the English speakers, the speakers were high proficiency, bilingual Spanish speakers. The Spanish speakers were all from the Gipuzkoa region of Spain, to ensure that the native accented condition was the native accent of the participants. There was no main effect of accent on sentence duration ($F(2,222) = 1.774$, $p = 0.172$). To minimize possible differences in prosody and speech rate, each speaker was asked to listen and repeat a sentence pronounced by a reference speaker (a non-native speaker of Spanish who first recorded all sentences with neutral intonation and slow speech rate). The sentences were presented blocked by accent and were split into 3 lists so that each participant would not listen to the same sentence twice and to ensure that different speakers were heard saying every sentence (25 sentences per each accent type, among which between 8 and 9 were produced by the same speaker). The order in which participants heard each accent was counterbalanced. Speech intelligibility measures were collected by having each participant repeat every sentence for each of the different accent conditions. This was done to ensure that any possible effects seen in the experiment were not due to reduced intelligibility or lack of attention. Sentences that were not repeated correctly by the participant were excluded from the analyses presented here (percent of sentences correctly repeated in each accent: native: 89%, dialectal: 88%, foreign: 80%[10].

4.3.3.5. Single Task-Baseline Two

A repetition of the first baseline task, consisting of 5 minutes of tracking the circle without their attention divided on any other task.

[10] see Appendix C for an overview of the selection process and a summary of the full dataset averages

4.3.3.6. Session 1: Working Memory N-Back Test

A visual N-back test was included both as a filler between the attention and memory tasks and because working memory has been shown to reliably predict people's ability to control their attention (for a review see: Fougnie, 2008) and has been found to be correlated with dual-task performance (for a summary of papers see: Gagné et al. 2017). Participants saw a fixation cross in the center of the screen followed by different colored squares that appeared for 500ms and then had 2 seconds to respond[11] whether the last square was the same color as the one N trials before. They completed 3 blocks of 50 trials: a 1-back block, a 2-back block and a 3-back block. Response times and accuracy (hit/miss) were collected. Working Memory scores were calculated by averaging correct responses with total trials per participant and thus creating one composite working memory score for the N-back performance of each participant (mean hit accuracy: 0.68, sd: 0.02).

4.3.3.7. Session 1 and Session 2: Old-New Recognition Memory Test

During both sessions the participants completed a visual Old (remember, know)-New task, using sentences that they had listened to and the new items that were created based on these sentences. Participants saw the sentence appear on the screen and were instructed to push a key response if they believed that they had listened to the sentence in the previous part of the experiment and a different key response if they believed that it was a sentence that they had not heard before. If the participant responded that they had listened to the sentence in the previous part they were asked to choose between 'remember' and 'know' responses. The difference between remember and know was explained in the instructions before beginning. The participants were told to respond 'remember (lo recuerdo)' when they explicitly remembered hearing that particular sentence in the previous part and they were told to respond 'know'(me resulta familiar) when they did not have a specific memory of listening to the sentence but it seemed familiar or they felt like they could have heard it. In case they reported having heard the sentence, they were then asked to indicate which accent they had heard the sentence in, having to choose between key responses for British, South American, or Spanish/Basque.

[11] Following a common N-back presentation format (e.g., Kane & Conway, 2007).

4.3.3.8. Visual Items for Old (Remember, Know) New Memory Test

During both sessions a visual memory test was performed, i.e., once after the attentional task and once 4-6 days later. The visual old-new memory test included all auditory items from the attentional portion of the experiments (old items) along with additionally created sentences (new items). These seventy-five unrecorded additional sentences were created by substituting two words in each original sentence with new words. All sentences were grammatically correct and plausible. New words were equated with their respective old word counterparts on the following lexical features, using the EsPAL Spanish database (Duchon et al, 2013): log frequency, number of letters, number of substitution neighbors, number of phonemes, number of syllables, number of phonological neighbors, familiarity, and imageability (all p values above 0.05). Items were equally selected across accent and speakers and divided into two subsets. Item subsets for the memory test were counterbalanced across sessions to ensure no repetition of items across sections[12]. These subsets were also equated on the abovementioned lexical properties (all p values above 0.05).

4.3.3.9. Session 2: Questionnaire and Self-Reported Listening Effort

On the date of return, participants were also told to fill out an online questionnaire [13]. Participants were given a questionnaire to determine to what extent they were exposed to English, Argentinean, and general foreign-accented speech in their daily life and also to acquire more information about their own accent, family, residence history, and subjective measures of accent-related differences (e.g. perceived effort on a scale of 1 to 10 of listening effort required by each accent).

[12] See Appendix D for list of old and new sentences

[13] Due to an initial error when marking the survey questions as obligatory, 5 questions were not answered by the first 5 participants. Thus, the survey results discussed in the following section only include 31 of the 36 final sample participants. See Appendix F, 2 for a list of survey questions.

4.4. Results

4.4.1. Attention

4.4.1.1. Listening Effort

After removing incorrectly repeated trials and converting the TOT scores into the listening effort equation used by Kemper et al, 2009 (Listening Effort = 100 × [(TOT Baseline[14] – TOT Dual Task) / TOT Baseline]), a two-way ANOVA with accent as a three-level factor (native, dialectal and foreign) and response type (listen and repeat) as a two-level factor was run on the listening effort data. A significant main effect of response type was found ($F(1,210)=4.12$, $p=0.04$), where listening was associated with higher effort than repeating (see Fig. 14). No significant main effect of accent was found ($F(2,210)=2.33$, $p=0.10$). Furthermore, no interaction was found between accent and response type ($F(2,210)=0.41$, $p=0.66$).

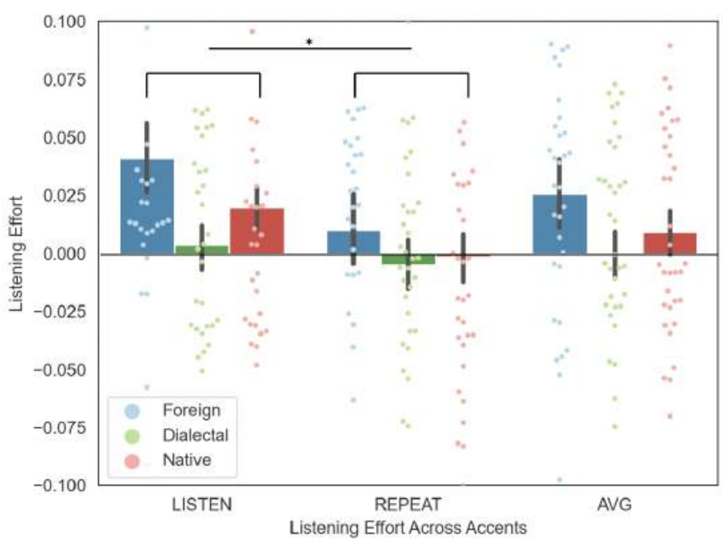


Fig. 14. Average listening effort scores separated by tracking when listening to, and repeating the sentences and the average of those two scores by accent, one standard error is shown.

.

To account for the considerable data variability, minimal dual-task effect, and a tentative indication of potential disparities in listening efforts among accents, supplementary exploratory t-test analyses were conducted on the average scores of the Listen and Repeat tasks (see Fig. 15). T-tests revealed a significant difference between dialectal and foreign accent, where foreign

[14] The baseline end condition was used here to account for the learning effects along the experimental session. Additional analyses with the averaged baseline (including the beginning and the end baseline conditions), can be found in Appendix E, 1.

accent was associated with greater listening effort than dialectal accent (t(35)=2.26, p=.03). Listening effort of foreign accent was also marginally higher than native accent (t(35)=1.85, p=.07). No significant difference was found between dialectal and native listening effort (t(35)=1.40, p=0.17).

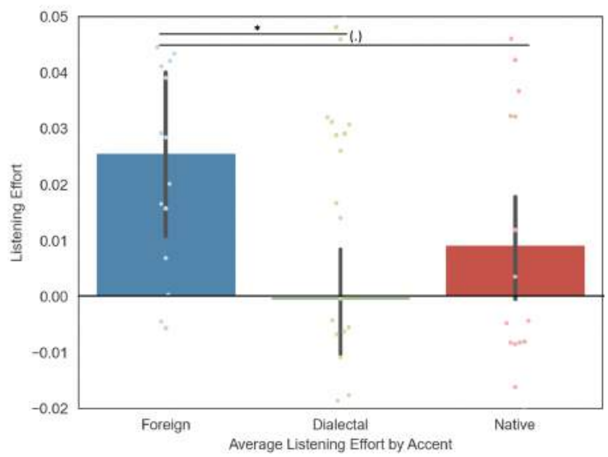


Fig. 15. Average listening effort scores (for listening and repeating) by accent, one standard error shown.

An additional exploratory analysis considered the tracking scores of all conditions (including baseline) to test for possible dual-task effects in the condition that requires the highest degree of listening effort (Listen). This was purely an exploratory analysis to understand the data better due to the lack of a clear dual-task effect in order to investigate whether there were any differences in tracking performance between listening to different accents while tracking and tracking alone (see Fig. 16). A one-way ANOVA with accent as a 4-level factor (native, dialectal, foreign, and baseline) was run. A marginally significant main effect of accent was found (F(3, 140)= 2.67, p=0.05). Follow-up linear model analysis revealed that there was a significant difference between the baseline and foreign TOT scores of the Listen condition (p=0.01). While there was no significant difference found between the baseline TOT score and the native (p=0.22) or dialectal TOT scores of the Listen condition(p=0.77).

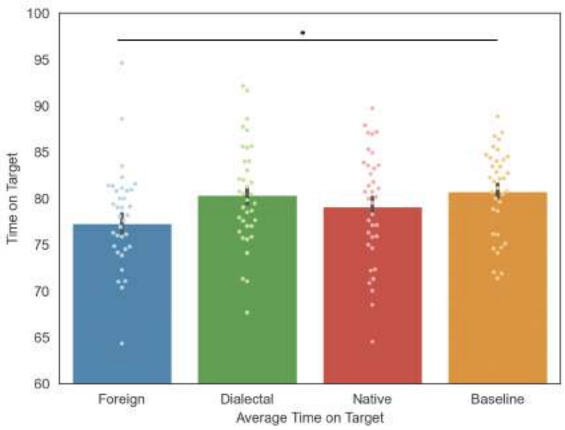


Fig. 16. Average time on target (TOT) during listening only scores during 3 dual-task conditions and single task baseline, one standard error shown.

4.4.1.2. Self-Reported Listening Effort

Self-reported listening effort scores from the post-test survey were also analyzed (see Fig. 17). On average, participants rated foreign accent as more difficult to pay attention to than native accent (foreign:4.7 effort out of 10, SD: 2.2, native:1.9, SD:1.9, dialectal:3.4, SD:1.9). An ANOVA was conducted on these self-reported measures including accent (foreign, dialectal, native) as a within-subject factor. Accent was found to have a significant effect on self-reported listening effort ($F(2,90)=13.76$, $p=0.0000006$), with foreign accent being perceived as the most difficult accent to pay attention to, followed by dialectal and then native (native vs foreign: $t(30)=5.61$, $p<0.0001$; native vs dialectal: $t(30)=3.97$, $p=0.0004$; foreign vs dialectal: ($t(30)=2.91$,$p=0.003$)[15].

[15] An ANOVA on the dual task listening effort measure of the same participant sample that completed the self-reported listening effort measures (n= 31) showed similar results to those reported in the main text. See Appendix E, 2.

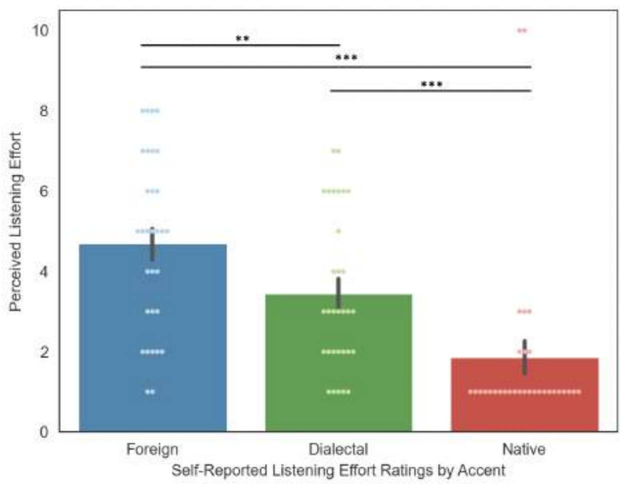


Fig. 17. Self-reported listening effort ratings of 31 participants (scale of 1 to 10), one standard error shown.

4.4.1.3. Is Listening Effort Related to Participant Background or Working Memory?

We wanted to test whether listening effort was affected by additional covariates of background collected by the questionnaire and working memory scores. The potential covariates were as follows: self-reported listening effort scores, accent type, interest in accent (scale of 1 to 10 how interesting/appealing was each accent to listen to), weekly social accent exposure, accent/history of parents, working memory score, whether they have lived or stayed in the UK or Argentina for an extended period, English level, whether they used headphones, and the calculated tracking speed of circle (rpm). Because of how many potential factors could be included, multiple models were first fitted using the ‘olsrr’ package in R (https://olsrr.rsquaredacademy.com/) and the best fitting one was determined based on the lowest AIC. Thus, the model determined to be best fitted was the only officially run, including only those factors. Our final model included the following factors: accent type, self-reported interest in the accents, whether they wore headphones and rpm of the circle. Only significant co-variates of the model are discussed (whether they wore headphones and the speed of the circle were not found to be significant factors). This model was run with both objective and subjective measures of listening effort. It was found that the participant’s self-reported interest in the accent has a significant effect on their objective attentional performance, namely that the more participants were interested in the accent of the secondary task, the harder the tracking during the primary task, and thus the higher the listening effort was (native: B = 0.05, SE = 0.02, p = 0.007; dialectal: B =0.05, SE = 0.02, p = 0.005; foreign: B = 0.05, SE = 0.02, p = 0.02).

Interest was also found to have an effect on self-reported listening effort. Participants' interest in either a foreign or dialectal accent positively affects their self-reported listening effort (foreign: B=2.0, SE=0.56, p=0.0005; dialectal: B=1.58, SE=0.53,p=0.004). However, this relationship disappears for native accent (B=-0.22, SE=0.55, p=0.70). Working memory also was related to only self-reported listening effort: lower working memory capacity was associated with higher reported listening effort overall (B=-27.81, SE=8.65, p=0.002).

4.4.1.4. Attentional Results Summary

Self-reported listening effort measures show a clear effect of accent, with foreign accent perceived as the most effortful one. Listening effort scores in the dual task only partially confirmed this pattern. While there was no clear accent effect, exploratory follow-up analyses provided some evidence for a higher listening effort of foreign accent compared to the other accents. In addition, we observed a dual-task effect (i.e. lower tracking performances in the dual task as compared to the single-task baseline) only when listening to the foreign accent. The listening effort scores were affected by the participants' accent interest, suggesting that the more interest participants reported for an accent (primary task) the more effortful the secondary task became.

4.4.2. Memory

4.4.2.1. Session 1 and Session 2 D' Scores

A two-way ANOVA with accent as a three-level factor (native, dialectal, foreign) and session as a two-level factor (session 1, session 2) was conducted on the calculated d' scores of the memory trials after removing all incorrectly repeated sentences from the attentional part of the experiment[16]. This was done to ensure equal intelligibility between accents, as trials which were repeated by the participant saying "I don't know" would not be encoded into memory, and

[16] The 'remember' 'know' classification of the memory trials was also analyzed but did not add anything to the general analysis. See Appendix F, 1 for details on this.

also to eliminate the possibility that the participants' own incorrect repetition of the sentence could interfere with their encoding of the sentence. As shown in Fig. 18, a significant effect of accent was found (F(2,210)=3.93, p=.02) and a significant effect of session was also found, where session 1 corresponded to higher d' scores overall than session 2 (F(1,210)=76.85, p=6.31e-16). There was no significant interaction between session and accent (F(2,210)=0.75, p=0.47).

To better understand the effects of accent on memory performance, the session 1 and session 2 d' scores were collapsed into one average score for each accent and a series of paired t-tests were performed. The results revealed that overall, participants performed significantly better on the foreign-accented trials (M=2.07, SD=0.67) compared to native (M=1.61, SD=0.88;t(35)=3.28, p=0.002). They also performed significantly better on dialectal accented trials(M=1.96 , SD=0.75) than native (t(35)= -2.25, p=0.03). No significant differences were found between the performance of foreign and dialectal trials (t(35)= 0.81, p=0.42).

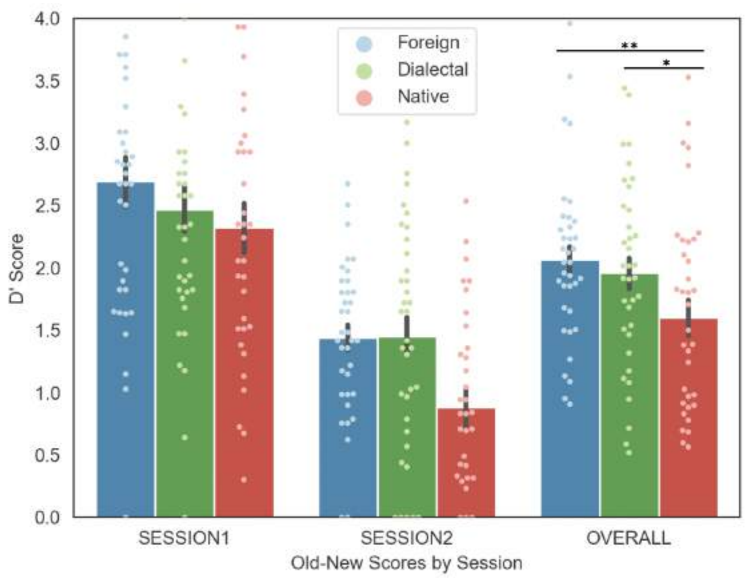


Fig. 18. Average d' scores for each of the three accented conditions of the old-new recognition memory test on session 1 and session 2 of testing, 4-6 days later, one standard error shown.

4.4.2.2. Are Memory Scores Related to Participant Background or Working Memory?

We wanted to test whether memory scores were affected by participants' questionnaire background (self-reported listening effort scores, accent type, interest in accent (scale of 1 to 10 how interesting/appealing was each accent to listen to), weekly social accent exposure, accent/history of parents, working memory score, whether they have lived or stayed in the UK or Argentina for an extended period, English level, whether they used headphones, and the calculated tracking speed of circle) and working memory. Because of how many potential factors could be included, multiple models were first fitted using the ‘olsrr’ package in R and the best fitting one was determined based on the lowest AIC. Thus, the model determined to be best fitted was the only officially run and only included the memory score average with the following factors: accent type, interest in accent, accent of parents, and working memory. Only significant co-variates of the model are discussed (working memory was a significant factor). The effect of foreign accent on memory performance described in the ANOVA was maintained ($B=0.38$, $SE = 0.17$, $p =0.02$) Self-reported interest in a particular accent was found to have an effect on memory score performance, where the performance of the native accented sentences in the memory task was lower the more participants reported to be interested in accents other than native (either foreign or dialectal; foreign: $B = -0.64$, $SE =0.21$, $p = 0.004$; dialectal: $B = -0.58$, $SE =0.21$, $p = 0.008$). Parents' accents were also found to have an effect on memory, where participants with both parents from the basque country (hence with full native accent exposure within the family) showed better memory performance on native accented trials ($B = 0.4$, $SE = 0.13$, $p = 0.004$).

4.4.2.3. Memory Results Summary

Overall, participants performed better on foreign accented trials and dialectal accented trials than native accented trials. This effect was modulated by participants’ interest in a specific accent: the more a participant reported interest in an accent other than native, the worse their memory performance on the native trials is. It was also influenced by a participant's home exposure to the native accent: if both parents of the participant had the native accent, they performed better on native trials.

4.4.3. Attention-Memory Relationship

A regression model was fitted to test whether the relationship between listening effort and memory performance was affected by accent. A significant negative relationship between attention and memory was present in native accent[17] (b =-5.67, SE =2.3, p =0.01), but this effect disappeared for foreign (b = -0.7, SE = 1.47, p = 0.63) and dialectal accent (b = 0.35, SE = 2.22, p = 0.87). Thus, when listening to native accent, higher listening effort was associated with lower memory performance. However, no such relationship exists when listening to either dialectal or foreign accents (see Fig. 19).

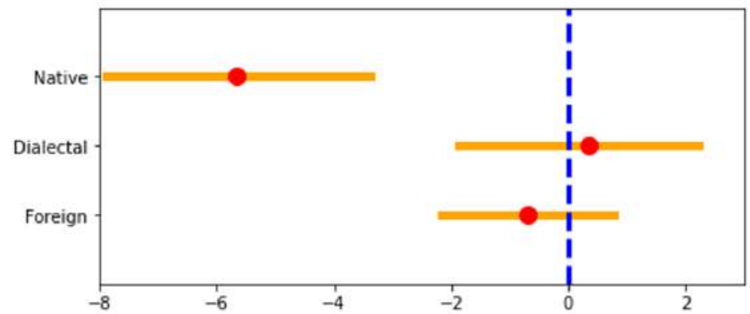


Fig. 19. Visual of the Beta coefficients of each accent.

4.5. Discussion

This study attempted to advance the debate on the attentional demands of variations in speaker accent. It further endeavored to elucidate the link between these attentional demands and memory by studying immediate and delayed information retention. Undoubtedly, the study of the effects of speaker identity on behavioral correlates of attention and information retention provides valuable contributions to the converging fields of memory, attention, and speech perception processing. All of these have been hitherto mainly studied independently.

Although previous literature has found that there were increased processing demands of foreign-accented speech (e.g. Clarke et al., 2004; Lev-Ari, 2014; Munro et al, 1995; Schmid et al., 1999, etc.), we were only able to find a clear effect of increased listening effort for foreign-accented speech in the subjective ratings of listening effort. In the listening effort measures collected from the dual-task paradigm, we observe weak evidence for differences between the listening effort of different accents. Critically, we did not observe a strong presence of a dual-task effect during our attentional task. Indeed, only in exploratory analyses did we observe a dual-task effect, with worse performance in the listening condition of foreign accent as

[17]A regression model was also fitted to test whether the relationship between self-reported listening effort and memory performance was affected by accent. No significant relationship was found (b = -0.05, SE = 0.05, p = 0.35).

compared to the baseline single-task measure. This is in contrast to other recent studies of listening effort that did find an effect of foreign accent on attentional demands (see Brown et al. 2020).

Interestingly, we found that listening to the sentences while tracking was significantly more effortful than repeating them while tracking. There are several possible reasons that could explain this finding. When participants engaged in the production task of repeating the sentences, it might have increased their overall level of engagement and focus on the task at hand. The active involvement and additional cognitive and motor demands of speech production could have heightened their attention and concentration, allowing for enhanced overall task engagement and performance in the circle tracking task. Alternatively, engaging in the speech production task while simultaneously tracking the circle might have triggered a sensorimotor coupling effect. The motor actions involved in speech production could have facilitated motor control and coordination, potentially leading to improved performance in the circle tracking task. Conversely, participants might have subconsciously prioritized the circle tracking task when repeating the sentences due to its immediate feedback nature. The real-time visual feedback provided by the circle tracking task might have led participants to allocate more attention and effort to it, resulting in improved performance. In contrast, when solely listening to the sentences, participants may not have perceived the same level of urgency or importance for the circle-tracking task.

While we did observe a trend toward higher listening effort for foreign-accented listening, the large variability of the data and lack of a dual-task effect made it difficult to understand the true demand on attentional resources. As mentioned, we supplemented the objective (dual-task) measure of listening effort with subjective effort ratings. Indeed, participants did show a strong effect of accent on subjective self-reported listening effort ratings. On average, participants clearly perceived the foreign accent as the most attentionally demanding, followed by dialectal and finally native. A few studies have argued that subjective measures of listening effort are reliable and even preferable measures of listening effort (see Peng and Wang 2019). Additionally, a recent study using Pupillometry, a dual-task, and a subjective rating of listening effort found a higher listening effort of non-native accented listening in both subjective and objective measures (Brown et al. 2020). However, overall, there have been limited studies that have attempted to link subjective ratings with objective task performance (see McLaughlin et al. 2021; Oosthuizen et al. 2020), and whether subjective ratings are a reliable measure of objective cognitive effort is still debated (Strand et al. 2018). If a subjective rating can be established as a reliable measure of listening effort, then listening to foreign-accented speech and to a lesser extent dialectal-accented speech can be said to be more attentionally demanding in this case. This finding, if reliable, would provide support for the Perceptual Distance Hypothesis, due to the gradient in perceived attentional effort across the accents.

However, in our study, the dual-task measure and subjective measure of listening effort do not seem to align. It is difficult to determine the cause of this because of the lack of significant effects in the dual-task paradigm. Perhaps they are truly not aligned measures, but alternatively, it could be that the dual-task paradigm used was simply not sensitive enough to reveal accent-driven trends from which to draw conclusions. In addition to a more reliable objective listening effort measure, it would be useful to have more information about the reliability of the subjective ratings in order to make stronger claims about listening effort. One encouraging finding regarding the possible reliability of the subjective listening effort measure is that working memory scores were found to be negatively correlated with self-reported listening effort, meaning higher listening effort ratings were associated with lower working memory capacity. Working memory has been shown to reliably predict people's ability to control their attention (for a review see Fougnie, 2008), suggesting that subjective listening effort could be a true reflection of attentional demand. Perhaps comparing objective and subjective listening effort using another objective measure, such as pupillometry, could clarify both the effects and whether the two measures are indeed related. Indeed, recent studies have successfully used pupillometry as a measure of cognitive effort (Brown et al. 2020; Winn et al. 2018; Miles et al. 2017; Borghini and Hazan 2018), showing promise for the technique.

Regarding memorization of information, several studies (Atkinson et al., 2005; Kavas et al., 2008; Lambert et al., 1960; Lev-Ari, 2012; Mayer et al., 2005; Munro et al., 1995; Ryan et al., 1982) have observed differences in comprehension or memorization of information delivered by different accented speakers. This has been postulated as being due to the decreased reliability, trustworthiness, and expectations of speech abilities that foreign-accented speech activates. Indeed, we did find an overall effect of accent on memory performance. Interestingly, memory performance was higher for non-standard (foreign and dialectal) than native accent. While numerically participants performed best on foreign-accented trials, this difference was not significant between foreign and dialectal. Furthermore, we did not find any interaction between session and accent, so we cannot conclude anything about whether we remember information received in certain accents longer than others. Perhaps with a larger sample size, whether there is a difference in memory performance between dialectal and foreign accents will become more clear. This would help clarify whether the memory data supports the Perceptual Distance Hypothesis or the Different Processes Hypothesis. While we do not see an attenuated enhanced memory performance for dialectal accent, due to the differentiation in performance between native and dialectal accent, the results could be argued to be in favor of the Perceptual Distance Hypothesis if the dialectal accent chosen had an especially large acoustic distance from the native accent. At the very least, the results do not seem to support the Different Processes Hypothesis due to the lack of differentiation between native and non-native accents.

While the results suggest that native accent listening is the least cognitively effortful and also associated with the lowest overall memory performance, the link between attentional load and memory needs to be further explored. We did find a negative relationship between listening effort and memory performance but only for native accent listening. This finding is in line with the Cognitive Load theory, where increased attentional demands lead to an overloading of the system that causes poorer encoding in the memory. Meaning, when listening to native-accented trials, the more listening effort a participant had to put into listening to any given trial, the worse their memory of the trial was. This link was not present for Foreign or Dialectal listening, where higher listening effort does not lead to poorer memory encoding, and in fact, results seem to show better memory encoding which is more closely aligned with the Cognitive Effort Theory. However, the memory performance of foreign and dialectal accents was not positively correlated with listening effort, meaning that we did not find statistical evidence for a link between attention and memory performances for foreign and dialectal accents. It appears that something unique happens when we listen to non-native or dialectal accents that enhances our memory. This effect exists regardless of whether we report interest in the accent and despite our own perception that foreign accent is the most difficult to remember (average difficulty ratings: foreign:5.3 out of 10, SD: 2.3, native:3.9, SD:2.6, dialectal:4.5, SD:2.3). Participants also reported foreign-accented sentences as the most difficult accent to repeat. (foreign:48%, native:0%, dialectal: 13%, none or I don’t know: 39%). If the Cognitive Effort theory is correct for non-native listening, this finding could make sense with the data. Perhaps this increased difficulty led to a narrowed focus that helped participants to remember the sentences better. However, because we did not observe an attention-memory relationship, further studies are needed to better understand how these two theories relate to accented speech comprehension. The finding that changing the accent also changes the relationship between attention and memory could potentially suggest that the Cognitive Effort and Cognitive Load theories are actually two sides of the same continuum. We speculate that it is possible for both to have a role in memory encoding, and it is simply that different circumstances can lead to different processing effects. Further research should attempt to disambiguate these potential circumstances. For instance, we know that listeners undergo automatic processing adjustments when listening to a foreign-accented speaker (Romero-Rivas et al. 2015). It is possible that processing adjustments affect how the allocation of our attentional resources affects our ability to memorize information, allowing us to mitigate or minimize the effects of a higher attentional load on our interactions. Future studies will hopefully continue to investigate the attention-memory link so that we can understand more fully the long-term effects of our various social information exchanges and interactions.

4.6. Conclusion

Accent strongly influenced self-reported effort measures, showing that foreign accent had the highest perceived effort reports. However, objective listening effort measures were only mildly affected, with only exploratory analyses showing the highest scores of listening effort for foreign accent. This seems to suggest that foreign accent might require the highest cognitive effort among the accents examined here, followed by dialectal accent (in line with the Perceptual Distance Hypothesis). Both foreign accent and dialectal accent were also associated with better memory performances as compared to native accent. While in native accent the higher the listening effort the lower the memory performance is (in line with the Cognitive Load Theory), both foreign and dialectal accent seem to disrupt this attention-memory trade-off. Specifically, when dealing with nonstandard accents, high levels of listening effort do not imply lower memory performances. Instead, foreign and dialectal accent have a beneficial effect on memory, which is at least partially in line with the Cognitive Effort Theory, while conclusions about the

Different Processes and Perceptual Distance Hypothesis are harder to make on memory performance. These results show that the relationship between attention and memory does not seem to be fixed but rather may change as a function of accent. The next chapter will extend the examination of foreign, dialectal, and native-accented speech to naturalistic extended speech.

5. Chapter V: A frequency investigation of accented discourse processing

5.1 Rationale

The final experimental chapter of this dissertation expands even further the level of linguistic analysis from the sentence to extended discourse. The chapter aimed to examine the effects of accent processing during continuous speech and also further addresses the competing hypotheses of dialectal accent processing: the Different Processes Hypothesis and the Perceptual Distance Hypothesis. We aimed to discover in which frequency bands native, dialectal and foreign speech processing differ and where dialectal accent is placed with respect to foreign accent in a naturalistic speech setting. While the first experimental chapter mainly explored linguistic challenges associated with accent variations from a bottom-up perspective and then second experimental examined the cognitive challenges of more lengthy speech in multiple accents, this chapter explores the interaction and integration of the linguistic and cognitive challenges that non-standard accents may impose on the listener in fluid conversation.

5.2. Introduction

As I have reiterated throughout the previous chapters, in our current society of global mobility, the study of speech accent is more relevant than ever. Many previous studies have shown impaired speech comprehension of unfamiliar or accented speech due to the unique challenges that non-native speech presents the listener (Floccia et al., 2006; Munro & Derwing, 1995b; Schmid & Yeni-Komshian, 1999; Anderson-Hsieh & Koehler, 1988; Major et al., 2002). As cross-cultural interactions become increasingly common, it is important to investigate the underlying mechanisms involved in processing non-standard speech in its multiple forms. Most studies focus on contrasting foreign accent with native accent using behavioral or evoked responses from electrophysiological (EEG) methods. Here we use a less common EEG technique to study how our brain responds to not only native and foreign accent but also dialectal accent.

While foreign accent is often the focus when discussing non-standard speech, dialectal speech can also be considered non-standard speech from the perspective of a native listener.

As discussed, a dialectal accent would come from a native speaker of the same language as the listener who is from a different county or geographical region. These accent types possess native phoneme variations, distinguishing them from foreign accents that often contain non-native variations resulting from the sounds or syntactic rules of the native language of the speaker. Because of these differences, this study aims to understand how dialectal accent is processed in relation to native and foreign accent.

It is generally agreed upon that foreign accent presents the greatest challenge to the listener as compared to native accent. Foreign accents are often identified more readily than dialectal accents and have a greater impairment on lexical retrieval (Girard et al., 2008; Floccia et al., 2009). However, the case of dialectal accent is still understudied and it is not clear what is its impact on brain responses relative to the foreign and the native accent. The findings of Girard et al., 2008 and Floccia et al., 2009 seem to be in line with the Different Processes, which suggests that there are qualitative differences in the processing mechanisms recruited for dialectal and foreign speech normalization (Floccia et al., 2009). In this case, the main qualitative distinction is between accents that share a native language and non-native accents. Another hypothesis, the Perceptual Distance Hypothesis, suggests that accents can be placed on a perceptual scale based on their acoustic distance from native speech and thus dialectal accent is processed more like an attenuated version of foreign accent (Clarke & Garrett, 2004). Behavioral studies have also provided support for this hypothesis, with Adank et al., 2009 reporting decreased processing speed as accents become less familiar or more foreign. An earlier study by Floccia et al., 2006 also provided evidence for this hypothesis when they found slower word recognition for words uttered in an unfamiliar dialectal accent, and even more slowly in a foreign accent.

In addition to behavioral techniques, EEG methods have been employed to investigate accented speech comprehension due to their ability to provide valuable time-sensitive information. One common technique for investigating the effects of accent on speech processing using electrophysiological methods is the Event-Related Potential (ERP) analysis. Event-related potentials (ERPs) are a type of neurophysiological measurement technique created by signal averaging used to study brain activity in response to specific events or stimuli. ERPs related to speech processing have shown that the time course of speech analysis changes as a function of accent

Early evoked responses typically associated with phonological analysis have mainly shown differences between foreign and native accent, supporting the Different Procceses Hypothesis (see the previous chapter, Thomas et al., 2022; Goslin et al., 2012). In late evoked responses the results are more mixed, with some studies supporting the Perceptual Distance Hypothesis (see Jiang et al., 2020) and others the Different Processes Hypothesis (see Hanulíková et al., 2012). While ERP studies have been instrumental in understanding differences between processing mechanisms of native and non-native accented speech, because this technique reduced the complexity of the electrophysiological signal to one time-dependent variable, it does not allow for tracking multiple rhythmic processes occurring in parallel. Therefore, it is also useful to focus on the frequency domain of the EEG signal when studying linguistic processes, especially in the case of sentences and discourses where the rhythmic activity of the brain can be tracked over large time windows.

While time-based approaches, such as ERP analysis, can provide us with beneficial information, time-frequency methods offer two main advantages. They provide information about electrophysiological activity in a naturalistic situation (passive listening to continuous speech) and they are able to detect electrophysiological brain activity undetectable by ERPs, reflecting different parallel oscillatory patterns that are not time-locked to the presentation of a stimulus. Oscillations are thought to reflect the coordinated activity of large groups of neurons and may play a critical role in synchronizing neural processing across different brain regions. These oscillatory patterns are generally categorized according to their frequency bands: delta (δ < 4 Hz), theta (θ 4-8 Hz), alpha (9-12 Hz), beta (13-25 Hz) and gamma (> 25 Hz). Power spectral density estimation is a time-frequency analysis that calculates the distribution of power in different frequency bands. Previous studies have linked different frequency bands to different aspects of speech processing, such as phonetic analysis, semantic processing, and syntactic integration.

In the field of linguistic research, three frequency bands have been characteristically linked to linguistic properties of speech: the delta, theta and gamma bands (for review, see Giraud & Poeppel, 2012; Grabot et al., 2017; Meyer et al., 2017). To a lesser extent, some research also has correlated alpha and beta band power with linguistic processes.

Within low-frequency ranges (i.e., δ and θ), it has been long suggested that theta oscillations reflect syllable tracking. Peña and Melloni found that theta power was higher listening to forward utterances as opposed to backward ones (which disrupt the syllabic structures of the utterance), suggesting that theta power is involved in tracking syllable patterns (Peña & Melloni, 2012). Several studies have provided evidence that theta oscillations synchronize with syllable onset (Luo & Poeppel, 2007; Howard & Poeppel, 2012; Peelle et al., 2013; Doelling et al., 2014) and thus aid in the identification of syllables. Limited additional studies have also suggested that theta oscillations may be additionally involved in lexical-semantic retrieval (Bastiaansen et al., 2008) and syntactic processing (Bastiaansen et al., 2002). Delta bands have been linked to intonation or prosodic processing due to their phase coherence with the pitch contour of speech (Giraud & Poeppel, 2012; Bourguignon et al., 2013; Mai et al., 2016).

Within high-frequency ranges (i.e., β and γ), more previous research has been done on gamma's role in speech processing. Traditionally, synchronization of the gamma band amplitude has been suggested to reflect phonemic-categorical perception (Lehongre et al., 2011). Ortiz-Mentilla and colleagues found that even as early as 6 months, high gamma power is enhanced during the discrimination of native phoneme contrasts, suggesting that it plays an early role in the perception and categorization of phonemic features, allowing for preferential processing of native phoneme contrasts (Ortiz-Mantilla et al., 2013). While low gamma synchronization has been linked to acoustic processing (Gross et al., 2013). The Beta frequency band has also been correlated with linguistic processes. It has been shown to decrease in situations of mismatch between semantic predictions and reality (Lewis & Bastiaansen, 2015; Lewis et al., 2016). Some scientists have also suggested that it may play a role in the processing of semantics (cf. Weiss & Mueller, 2003). The mid-range alpha frequency, however, is mostly attributed to more general cognitive functions and has been classically related to the inhibition of task-irrelevant information (for review, see Klimesch et al., 2007). Despite alpha being associated most prominently with generalized top-down mechanisms, some previous works have provided evidence that it may play a role relevant to sentence comprehension in verbal working memory (Krause et al., 1996; Maltseva et al., 2000; Jensen et al., 2002; Sauseng et al., 2005); Leiberg et al., 2006).

While there are many studies linking various neural oscillations to linguistics processes, very little work has been done to differentiate dialectal and foreign accents in the frequency domain. While behavioral and ERP studies have provided evidence for both previously proposed hypotheses, this is the first study comparing neural oscillations of foreign, dialectal and native accents. Examining discourse through the frequency domain allows us to see how speech accent affects brain rhythmic activities at different frequency ranges thus enabling us to understand the robustness of previously observed evidence in favor of either the Perceptual Distance or Different Processes Hypothesis.

5.2.1. The Present Study

The present study aims to investigate how the brain processes accented speech across different frequency bands using electrophysiological methods. We critically evaluate the aforementioned previously proposed hypotheses of accented speech processing in order to clarify the processing of foreign, dialectal and native accented speech. Special attention is placed on disentangling how we process dialectal accent from both native and foreign accent during the listening of short stories. Participants also performed a target detection task and answered comprehension questions. Analyses include Power spectral density estimation for each accent condition because of its relevance to EEG analysis of extended speech. Specifically, we focus on Delta, (1-3 Hz) Theta (4-8 Hz) and Low Gamma (25-35 Hz) waves, associated with prosodic, syllabic and phonemic processing, respectively (see Meyer, L., 2017 for a review).

According to the Perceptual Distance Hypothesis, we should observe a gradient in the average power amplitude of frequency bands that are implicated in the processing of native, dialectal and foreign accent, respectively. Conversely, results showing that dialectal and native accents are processed similarly would support the Different Processes Hypothesis. Thus, we would see a categorical discrimination between the foreign accent and the other more standard accents (native and dialectal).

5.3. Materials & Methods

5.3.1 Participants

Thirty Spanish natives[18] participated in the study. One participant was excluded due to low performance on the target detection task (mean accuracy: 37.5%). Another participant was excluded for low comprehension question performance (mean accuracy: 53.3%). The final sample of participants consisted of 28 females[19] (mean age: 22.7 years, SD: 3.56, age range: 19-31 years, Spanish age of acquisition: 0). All participants lived in the Basque country and considered the Basque-Spanish accent as their native accent. All participants were right-handed and had normal or corrected-to-normal hearing and vision. No participant reported a history of neurological disorders. All participants signed an informed consent form before taking part in the study that was approved by the Basque Center on Cognition, Brain and Language ethics committee. They received monetary compensation for their participation.

5.3.2. Materials

Three dialogues were recorded by six female speakers in their thirties with differing accents. Each dialogue was recorded by two speakers of the same accent (native, foreign, dialectal). Each pair of same-accent speakers recorded the three stories (total of recorded stories: 9). Story type and pair of speakers were fully counterbalanced so that each participant listened to dialogues produced by each pair of speakers without repetitions (total of 3 stories presented). They were recorded in a Basque Spanish native accent, a Cuban Spanish accent and an Italian Spanish accent (i.e., hereafter native, dialectal and foreign accent, respectively). The native speakers were born and lived in Spain (Basque country). The dialectal and foreign speakers were chosen for their strong accents and high level of Spanish (in the case of the Italian speakers). Overall, the recordings did not differ in duration (ms) across accents [foreign: 519.7, SD:128.7; native: 505.7, SD: 137.2; dialectal: 519.9, SD:131.6; one way ANOVA: ($F(2,117)=0.15$, $p=0.86$)]. Accent strength ratings were collected from a separate normative study consisting of eleven participants (average age =23, SD=10) who completed a short online survey where they listened to clips of each accent and rated the accent strength from 1 (mild accent) to 5 (strong accent), showed a clear effect of accent (one way ANOVA: $F(2,10)=6.01$, $p=0.009$). Follow-up

[18] Due to their geographical location in the Basque Country, all participants were also fluent in Basque as early L2

[19] Only females were recruited for this study in order to avoid cross-gender listening effects (Banaji and Hardin, 1996; Cacciari and Padovani, 2007)

analyses of the clip ratings corrected with the Fisher's Least Significant Difference (LSD) showed that the dialectal accent was marginally significantly different than the native accent ($t(10)=2.13$, $p=.06$), the foreign accent was significantly different from the native ($t(10)=3.16, p=.00$) and the dialectal and foreign accents were not rated significantly different from each other in terms of strength ($t(10)=1.15$, $p=.28$). 10 comprehension questions about the dialogues were also created[20].

5.3.3. Procedure

Participants carried out a blocked-design experiment consisting of 4 blocks, three accented blocks and one silence block. Participants were seated in a sound-attenuated room in front of a computer screen and were asked to listen to narrative dialogues and occasionally perform a target detection task to maintain the listener's attention by pressing the spacebar with both index fingers when they heard certain target words ("amigo/a/os/as"). After each block, they filled out a questionnaire with 10 comprehension questions about the dialogue. A fixation cross remained on the screen for the entirety of each block. Each block consisted of a dialogue (or silence) presented through speakers while a fixation cross was displayed on the screen throughout the entirety of the dialogue. The experimental session lasted about an hour. During the silence block, participants still occasionally heard the target words and had to press the spacebar. The block accent order was counterbalanced across participants.

5.3.4. EEG Data Recording and Time-Frequency Analyses

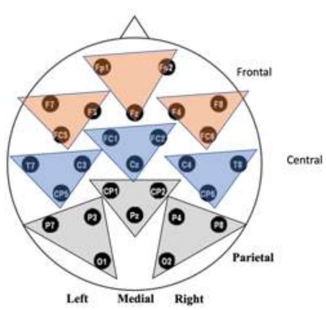


Fig. 20. Schematic of electrode montage with topographic organization labeled.

[20] See Appendix G for list of comprehension questions

The EEG signal was recorded from 27 channels placed in an elastic cap: Fp1, Fp2, F7, F8, F3, F4, FC5, FC6, FC1, FC2, T7, T8, C3, C4, CP5, CP6, CP1, CP2, P3, P4, P7, P8, O1, O2, Fz, Cz, Pz (see Fig. 20). Two external electrodes were placed on the mastoids, two were on the ocular canthi, one above and one below the right eye. All sites were referenced online to the left mastoid. Data were recorded and amplified at a sampling rate of 250 Hz. Impedance was kept below 10 KΩ for the external channels and below 5 KΩ for the electrodes on the scalp. EEG data were re-referenced offline to the average activity of the left and right mastoid. A low-pass filter of 30 Hz and a high-pass filter of 0.01 Hz were applied. Vertical and horizontal eye movements were corrected by performing an Independent Components Analysis (ICA). The fastICA method was used. Time-frequency analysis of continuous EEG data was done with a Morlet wavelet decomposition using MNE software (Gramfort et al., 2013). This method was used to decompose trial time-frequency values between 1 and 35 Hz for the 28 electrodes placed on the scalp (steps = 13). The average total power values were baseline corrected with a log-ratio baseline correction (-500 to 0 ms). For each accent condition and frequency range, the resulting power was averaged across time and channels.

5.4. Results

5.4.1. Behavioral Results

Participants showed relatively high accuracy during the online target detection task meant to test attention (mean overall accuracy: 71.4%, SD:13.2). Participants also showed high accuracy on comprehension question performance (mean overall accuracy:83.6%, SD:8.7).

5.4.2. Time-Frequency Results

Power amplitude modulation in the EEG of participants was analyzed with a mixed linear model with participants as random intercept and accent as fixed factor (native, dialectal and foreign). An accent effect was observed only within the gamma frequency range (see Fig. 21).

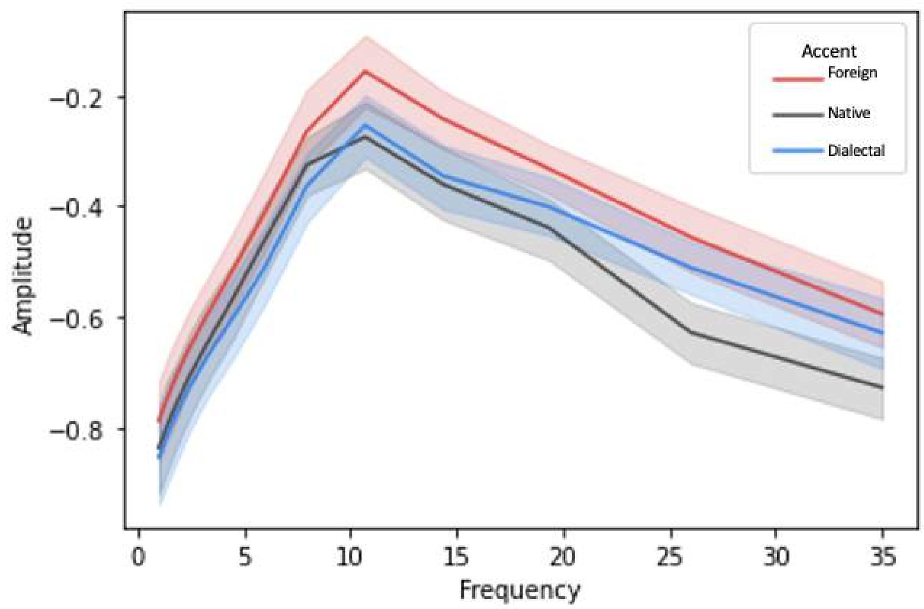


Fig. 21.. Power amplitude modulation in the EEG across accents.

Delta (1 to 3 hz) prosody

In delta frequencies we did not observe any power difference between the three accents (native vs. dialectal: β=-0.02, *SE*=0.09, *t*=0.19, *p*=0.84; native vs. foreign: β=0.05, *SE*=0.09, *t*=0.53, *p*=0.60; foreign vs. dialectal: β=-0.07, *SE*=0.09, *t*=0.72, *p*=0.47), see Fig. 22.

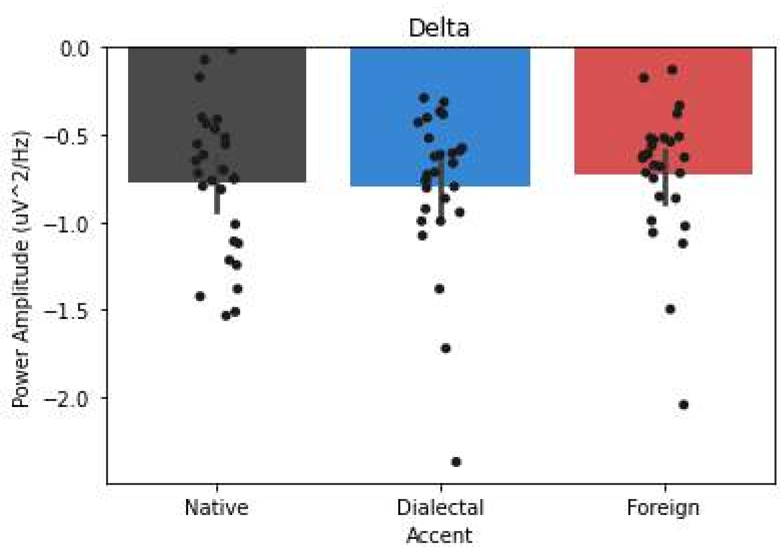


Fig. 22. Average delta power between accents, ci=95.

Theta (4 to 8 hz) syllable

In theta frequencies we did not observe any power difference between the three accents (native vs. dialectal: β=-0.04, *SE*=0.08, *t*=0.66, *p*=0.62; native vs. foreign: β=0.05, *SE*=0.08, *t*=0.49, *p*=0.51; foreign vs. dialectal: β=-0.09, *SE*=0.08, *t*=1.15, *p*=0.25), see Fig. 23.

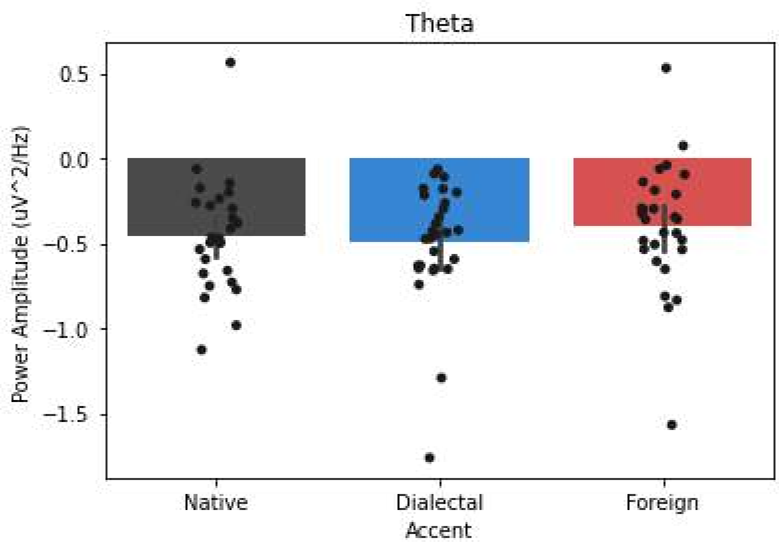


Fig. 23. Average theta power between accents, ci=95.

Alpha (9 to 12 hz)

In alpha frequencies we did not observe any power difference between the three accents (native vs. dialectal: β=0.02, *SE*=0.08, *t*=0.26, *p*=0.79; native vs. foreign: β=0.12, *SE*=0.08, *t*=1.48, *p*=0.14; foreign vs. dialectal: β=-0.10, *SE*=0.08, *t*=1.21, *p*=0.22), see Fig. 24.

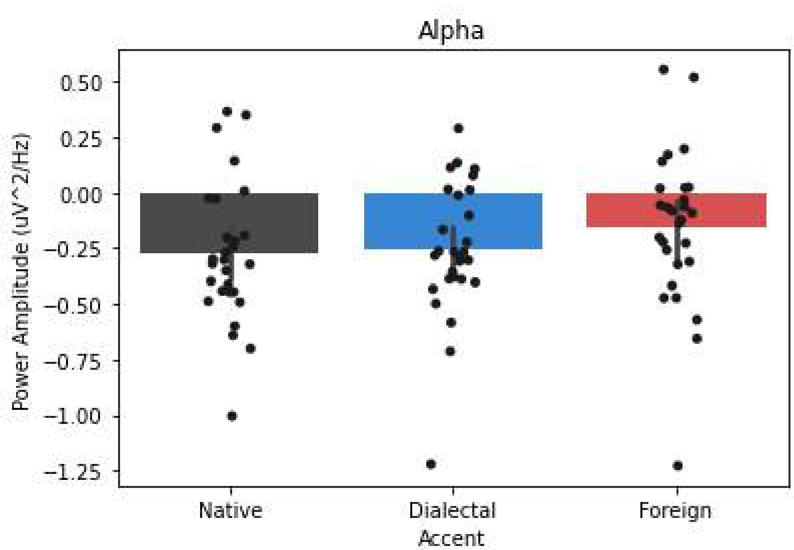


Fig. 24. Average Alpha power between accents, ci=95.

Beta (13 to 24 Hz)

In beta frequencies we did not observe any power difference between the three accents (native vs. dialectal: β=0.03, *SE*=0.07, *t*=0.39, *p*=0.69; native vs. foreign: β=0.11, *SE*=0.07, *t*=1.66, *p*=0.10; foreign vs. dialectal: β=-0.08, *SE*=0.07, *t*=1.27, *p*=0.21), Fig. 25.

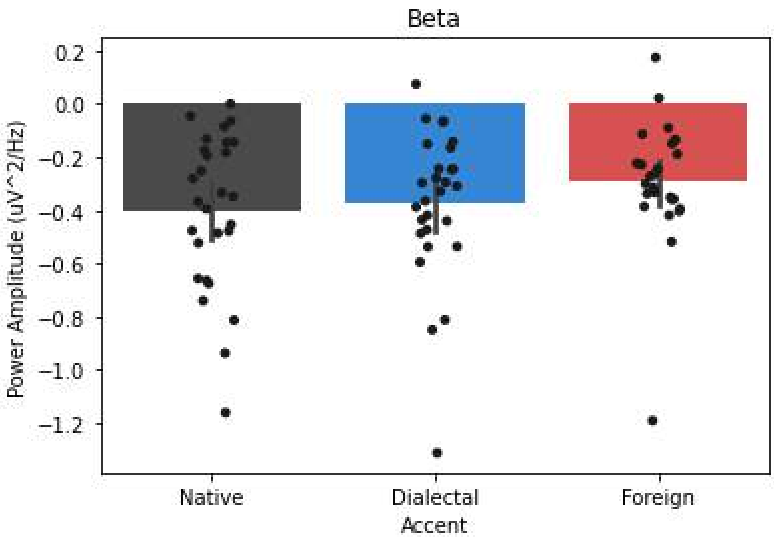


Fig. 25. Average Beta power between accents, ci=95.

Gamma (25 to 35 hz) phoneme

In gamma frequencies we found a marginally significant difference between native and foreign accents (native vs. foreign: β=0.15, *SE*=0.08, *t*=1.97, *p*=0.048). However, no difference was found between dialectal accent and either native or foreign accent (native vs. dialectal: β=0.11, *SE*=0.08, *t*=1.40, *p*=0.16; foreign vs. dialectal: β=-0.04 *SE*=0.08, *t*=0.58, *p*=0.56), see Fig. 26.

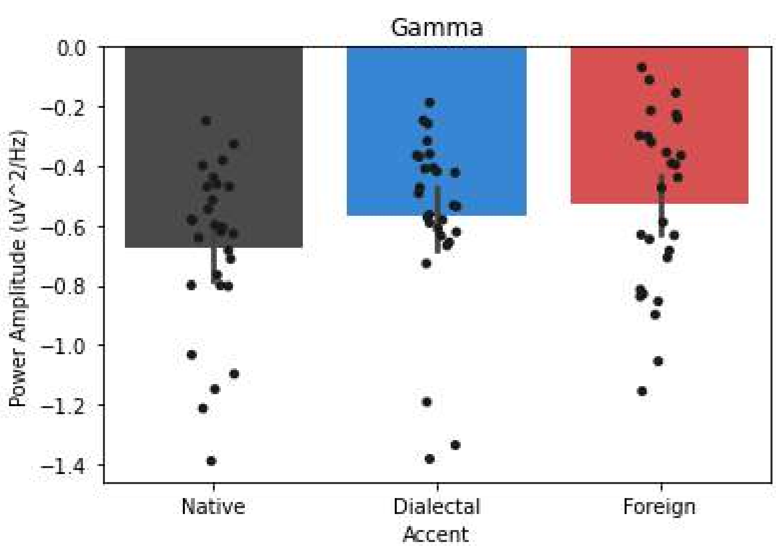


Fig. 26. Average Gamma power between accents, ci=95.

5.5. Discussion

This study aimed to advance previous investigations of online accented speech processing through the lens of oscillatory activity power amplitude fluctuations during listening to different speech accents. We further aimed to better understand the similarities and/or differences between mechanisms of dialectal and foreign-accented speech processing by examining the results in the context of two hypotheses of accented speech processing: the Different Processes Hypothesis and the Perceptual Distance Hypothesis.

Previous studies have emphasized processing differences between native and foreign speech processing (Lane, 1963; Lev-Ari & Keysar, 2012; Lev-Ari, 2017; Munro & Derwing, 1995b; van Wijngaarden, 2001). But whether the native listener processes native category variations, i.e. dialectal accent, according to acoustic distance from native speech (Perceptual Distance Hypothesis) or with 'nativeness' as a processing category (Different Processes Hypothesis) is less clear.

This study examined these hypotheses by investigating how the brain processes accented speech across different frequency bands using electrophysiological methods. We focused on Delta, (1-3 Hz) Theta (4-8 Hz) and Low Gamma (25-35 Hz) waves, because of their previous association with prosodic, syllabic and phonemic processing, respectively. We also considered Alpha and Beta because of their relation to attentional processing and top-down mechanisms.

We only found a significant difference between the power amplitude of native and foreign accent in the low gamma frequency range. This finding shows that unique processing mechanisms are required for the phonological processing of foreign accent throughout discourse processing. This provides evidence for the Different Processes Hypothesis by supporting differential processing for non-native accents but not for dialectal accent. Indeed, an exploratory follow-up Bayesian analysis provided a slight preference for the null hypothesis (BF: 0.89) and thus additional evidence that dialectal accent phoneme processing is similar to native accent phoneme processing, perhaps due to the 'nativeness' of the deviations in pronunciation, meaning that 'coherent deviations' (see Wells, 1982; Goslin et al., 2012) of the speech in a dialectal accent are easily adapted to by listeners while foreign deviations make acoustic extraction more difficult.

These findings seem to somewhat align with previous ERP studies of accent processing, including our previous study of single word processing, see Chapter 3 (Thomas et al., 2022 but see Goslin et al., 2012) where we found a difference between foreign and both native accent variations at early components associated with the extraction of acoustic features. Similarly, at components associated with lexico-semantic integration, we no longer observed accent differences.

The results of this study have similarly suggested that non-native accent affects early stages of phoneme processing. They have further suggested that these effects are uniquely seen for non-native accents rather than unfamiliar native variations. These results contribute to the literature supporting the Different Processes hypothesis and further support the idea of a binary native/non-native processing mechanism in the more naturalistic setting of extended discourse. Furthermore, accent seems to have the largest effect on phonological analysis while successfully resolved in the semantic phase.

Previous studies show that listeners employ top-down resources to process speech in challenging situations, engaging contextual cues, predictive processing, and prior knowledge about the topic to enhance comprehension (Dave et al., 2021; Foucart et al., 2015; cf. Schiller et al., 2020). However, in the alpha frequency band, associated with top-down mechanisms and attentional control/inhibition, we did not observe any significant effects. This is surprising in light of the previous chapter's findings, as we may expect to see some engagement of inhibitory processing in attentionally demanding situations, especially given the lingering effects of both dialectal and foreign accent-delivered speech on memory. Thus, whether accent is truly resolved at the phonological level or simply mitigated by the engagement of top-down strategies remains to be further explored and should be additionally scrutinized in future studies.

5.6. Conclusion

We found that while in higher frequency ranges foreign accent processing is differentiated from power amplitudes of native accent processing, in low frequencies we do not see any accent-related power amplitude modulations. This suggests that there may be a difference in phoneme processing for native and foreign accent, while we speculate that top-down mechanisms during discourse processing may mitigate other effects observed with shorter units of speech.

6. Chapter VI: General Discussion

In this chapter, I revisit the research questions laid out in section 2.2.1. and address them in light of the main findings of the studies presented in this doctoral work. I first summarize the results from each study. I then discuss the theoretical implications of these findings and situate them in the larger context of the general framework of language processing in light of the contrastive results on the effects of different types of accent on the time course of language analysis. The present doctoral work has primarily aimed to explore the linguistic and cognitive challenges of accent processing and thus contribute to the knowledge about the following remaining questions in the field of accented speech processing:

RQ1. How does accent affect early processing mechanisms during isolated single-word processing?

RQ2. How does accent affect the listening effort and memory of sentence processing?

RQ3. How does accent affect processing mechanisms during extended discourse processing?

The work attempted to explore these questions in the context of the challenges they present to the native listener. The linguistic challenges were the focus of RQ1, where I have explored early processing mechanisms of accent at the phonological and lexical-semantic level of single-word processing. RQ2 was focused on the cognitive challenges of accented speech processing, exploring the attentional listening effort and memory effects of varying the accent as the linguistic level of processing becomes slightly more complex. Finally, RQ3 explored the challenges related to the integration of cognitive and linguistic processing due to the interaction of bottom-up and top-down mechanisms that the longer time course of fluid discourse allows for. Below, I will discuss how the studies carried out for this work address these outstanding questions.

6.1. Summary of findings and general conclusions

6.1.1. Linguistic challenges: effect of accent on phonological and lexical-semantic processing in isolated words

The first study aimed to disentangle the competing hypotheses of dialectal accent processing: the Different Processes Hypothesis and the Perceptual Distance Hypothesis, and understand early accent processing mechanisms. If we recall from previous chapters, the Different Processes Hypothesis posits that the mechanisms of foreign and dialectal accent processing are qualitatively different (Floccia et al., 2009), while the Perceptual Distance Hypothesis posits that the mechanisms underlying dialectal accent processing are attenuated versions of those of foreign (Clarke & Garrett, 2004). Electrophysiological methods were employed due to their ability to provide valuable time-sensitive information. Event-Related Potential mean amplitudes were extracted: P2 [150-250 ms], PMN [250-400 ms] and N400 [400-600 ms]. Support for the Different Processes Hypothesis was found in different time windows. Results show that early processing mechanisms distinguish only between native and non-native speech, with a reduced P2 amplitude for foreign accent processing, supporting the Different Processes Hypothesis. Furthermore, later processing mechanisms show a similar binary difference in the processing of the accents, with a larger PMN negativity elicited in the foreign accent than the others, further supporting the Different Processes Hypothesis. Our results support the idea that, even during mostly passive isolated-word listening, indexical features, such as accent, affect our processing at early stages. Furthermore, our findings support that foreign accents are processed differently from dialectal or native accents, requiring more effort to extract acoustic characteristics and to normalize them into expected phonological representations. Despite this, it appears that we are able to successfully normalize non-native speech and thus by later processing stages accent no longer affects our processing or ability to integrate speech sounds into lexical information (no N400 modulation by accent).

These findings fit into the framework of cascading effects, suggesting that phonological deviations can cause processing difficulties that can cascade into later levels of linguistic processing if not successfully normalized. Because of this normalization, the results align with the Perceptual Assimilation Model, which suggests that during early stages of speech perception, listeners categorize speech sounds based on their native language phonetic inventory. PAM proposes that listeners assimilate non-native speech sounds into categories that exist in their native language, leading to potentially reduced processing for non-native sounds. The reduced P2 amplitude for foreign accent processing supports the idea that early processing mechanisms differentiate between native and non-native speech. The findings also indicate that, despite initial differences in processing, non-native speech can be successfully normalized and integrated into lexical information during later processing stages. This supports the idea that, over time, listeners become more proficient at extracting relevant acoustic characteristics and adapting to non-native speech patterns, as proposed by the Speech Learning Model. The SLM suggests that with exposure and practice, listeners can overcome difficulties associated with foreign accents and integrate speech sounds into their established phonological representations. This appears to be the case even at the level of a single word, where we have seen that by 400 ms post-word onset, the differences between different accents become null.

6.1.2. Cognitive challenges: accent effect on listening effort and memory processing in sentences

Study 2 examined the effects of accent on listening effort, on memory and on their interaction. Recall that, based on the gap in this literature, three different models, within the domain of the allocation of attentional resources and subsequent effects on memory retention, have been proposed: the Cognitive Load Theory, the Cognitive Effort Theory and the Indexical Free Theory. Support of the Cognitive Load Theory would predict that as a result of higher attentional demands, there would be poorer information load and memory retention performance for foreign compared to native accents. Conversely, Cognitive Effort Theory would predict that higher attentional demands would result in better information load and memory retention performance for foreign compared to native accents. Finally, the Indexical-Free Theory would predict that regardless of varying attentional demands, there will be no difference in the memory retention performance for foreign compared to native accents. Using the dual-task paradigm to measure attention through listening effort, we hypothesized that deviations from a native accent in a listening task would increase demands on attentional resource allocation (listening effort), and thus decrease performance on the secondary task. To investigate these theories, Spanish-Basque early bilinguals from the Basque country followed a dual-task paradigm, listening to Spanish sentences in multiple accents (native, dialectal, foreign) in a blocked design while completing a secondary task (circle tracking on the screen). For each accent, listening effort measures were collected together with memory scores in an immediate and in a secondary session (4-6 days). Results show that while higher listening effort is associated with lower memory performance for native accent, foreign and dialectal accent seem to disrupt this attention-memory trade-off. In fact, foreign and dialectal accent appear to have a beneficial effect on overall memory performance independently of listening effort. Results suggest that the relationship between listening effort and memory does not seem to be fixed but rather may change as a function of accent.

It is not entirely clear whether these findings support the Cognitive Effort or Cognitive Load Theory, as support could be argued for both theories. The finding of a negative relationship between listening effort and memory performance of native accented listening seems to support the Cognitive Load Theory, due to the inverse relationship between attentional demands and memory performance. However, the enhanced memory performance of foreign and dialectal trials is more aligned with the Cognitive Effort Theory, because despite participants perceiving these speech variations as more attentionally demanding, they performed better. It is important to note that no statistical evidence was found for a link between attention and memory in non-native accents making it less clear whether there are interfering mitigating factors in non–standard accent processing. It is clear, however, that these findings do not support the Indexical Free Theory. Indeed, both foreign and dialectal accents seem to have an impact on memory performance. If we think also about the theories of dialectal accent processing, again we seem to have difficulties discerning which theory is supported. Because there is a difference between dialectal and native, we might support the Perceptual Distance Hypothesis. However, because there is no difference between foreign and dialectal accent, the support is less clear. What is clear is that something unique appears to be happening for both dialectal and foreign accent when encoding the speech into a recallable memory representation. This supports the Exemplar model of speech perception, which posits that utterances are stored in the mind as distinct exemplars, retaining detailed acoustic and phonetic information rather than as abstract representations without indexical information. More evidence for exemplar models is found when we consider that participants consistently rated foreign accent as the most difficult accent for them to remember. Despite the fact that the experimental evidence shows that it was in fact the opposite, participants still felt non-native accented trials gave a salient impression. Perhaps participants mistook how foreign and dialectal accents stuck out in their memory for difficulty. Future studies should continue to explore the link between attention and memory in accented speech processing to better understand what mitigating factors may be integral to the effects observed in the recognition memory performance.

6.1.3. Cognitive-linguistic integration: effects of accent on extended discourse processing

Study 3 also aimed to examine the effects of accent processing during continuous speech and also further addresses the aforementioned competing hypotheses of dialectal accent processing: the Different Processes Hypothesis and the Perceptual Distance Hypothesis. Electroencephalographic data was collected from participants listening to 3 Spanish dialogues of 6 minutes recorded in different accents (native, dialectal, foreign). Data Analyses included Power spectral density estimation (1-35 Hz) for each accent condition. We aimed to discover in which frequency bands native, dialectal and foreign speech processing differ and where dialectal accent is placed with respect to foreign accent. Results show power differences in the Gamma frequency range. While in higher frequency ranges foreign accent processing is differentiated from power amplitudes of native accent processing, in low frequencies we do not see any accent-related power amplitude modulations. This suggests that there may be a difference in phoneme processing for native and foreign accent, while we speculate that top-down mechanisms during discourse processing may mitigate the effects observed with short units of speech.

According to the Standard Two-Step Model of Language Interpretation, the power differences in the Gamma frequency range could be associated with the initial perceptual processing stage, where acoustic information is transformed into neural representations. In the context of the Standard Two-Step Model, this could imply that foreign accents introduce novel or unfamiliar phonetic patterns that require additional neural processing to recognized them, as compared to native accents. Indeed, this appears to be what we see in study 1, with accent modulating the P2 component reflecting acoustic processing. However, the absence of accent-related power amplitude modulations in low frequencies associated with lexical and semantic processing suggests that as the level of linguistic analysis moves away from acoustic signal processing to incorporating more contextual cues, the effects of accent are minimized. This is observed at the level of the single word as well. In study 1, after the phonological normalization, we no longer observed any ERP modulations within the component associated with auditory semantic processing and lexical-semantic integration. The differences observed between accents in both EEG studies provide evidence for The Different Processes Hypothesis due to the nativeness of the speaker seeming to influence the processing of the listener more than the familiarity of accent.

This study provides an optimistic perspective on accent processing in a naturalistic setting. While bottom-up processing differentiates between accents, this differentiation does not seem to have long-standing effects that carry over into higher-level processing. Perhaps with interactions with substantially impaired intelligibility or a high number of semantic or syntactic errors due to the non-native status of the speaker, we would continue to observe effects across frequency bands. However, in most extended conversations with accented speakers, minimal errors or linguistic expectation violations seem to be rapidly dealt with by engaging top-down resources like our broader cognitive system and knowledge about the broader context of the conversation and therefore reduce or eliminate potential communicative disruptions due to accent.

6.2. Broader theoretical implications

The experiments presented in this dissertation aimed to further disentangle a variety of hypotheses regarding accent processing and the broader link between attention and memory processes. In addition to testing these hypotheses, it is important to place the work in the context of the theoretical frameworks discussed in the introduction.

As previously explained, the Dual-Stream Model and the TRACE model represent prominent frameworks for understanding speech perception from a neurobiological perspective. While the Dual-Stream Model and the TRACE model offer different perspectives on speech perception, they are not mutually exclusive. In fact, they provide complementary insights into different aspects of the perceptual and cognitive processes involved. The Dual-Stream Model emphasizes the neural pathways and processing streams involved in analyzing acoustic information and integrating contextual cues, while the TRACE model focuses on the interaction of phonemic, lexical, and semantic representations. Together, these models contribute to a more comprehensive understanding of how speech is perceived and processed in the human brain. In the context of speech accent processing, both models should be considered. The combined results of ourexperiments do support the plausibility of separate processing streams for acoustic and lexical-semantic processing. There certainly seems to be online differentiation between accents mostly for acoustic processing over higher-level processing. However, the complex interaction between levels of linguistic analysis, and the effects of accent on even general cognitive processes like attention and memory suggests that there may also be a multidirectional connective web between different processes.

Future research should continue to refine and expand these models, incorporating new findings from neuroimaging, computational modeling, and behavioral studies to further unravel the complexities of speech perception especially in diverse interactions. The Dual-Stream Model posits that there are parallel routes in the mapping from acoustic input to lexical phonological representations (see Fig. 27). Our studies provide confirmatory evidence for this hypothesis for accent processing, as in study 3 we observed effects in gamma oscillations corresponding to the pathway that samples acoustic input at a fast rate (gamma range). This is related to resolving segment-level information as opposed to the other pathway appropriate for resolving syllable leveling information. The existence of the two pathways helps explain why we only see modulations in the gamma range and acoustic extraction not in frequency ranges associated with syllable processing..

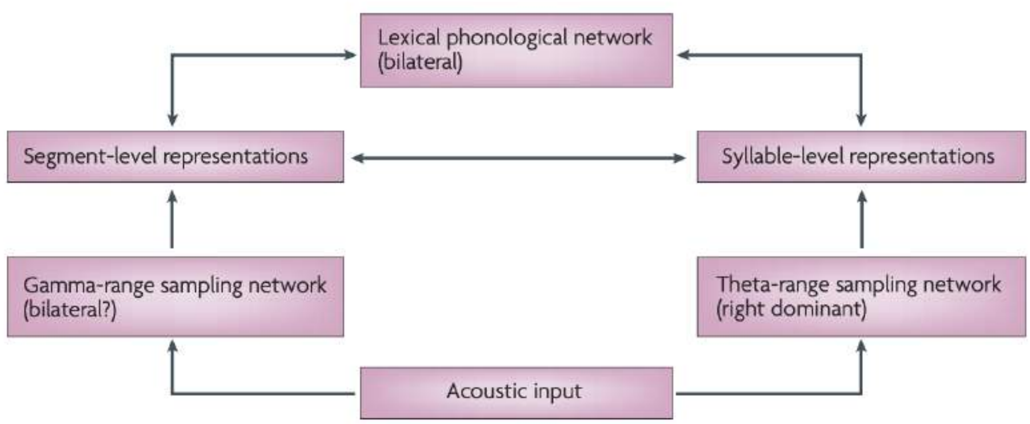


Fig. 27. Parallel routes of mapping acoustic input to lexical phonological representations from the dual-stream model. Figure taken from Hickok & Poeppel, 2007.

Indeed, further contextualizing the findings of this dissertation in the Dual-Stream Model, we can say that accent seems to have a modulatory effect, especially in the phonological step of speech perception. That is, in the part of the phonological network in the ventral stream (see Fig. 28). The phonological processing area as defined by the Dual-Stream Model, appears to be localized around the Superior Temporal Gyrus (STG), perhaps this model can be expanded to include a functional pathway between this area and the hippocampus, which is highly involved in memory formation (Ranganath et al., 2005). While not directly connected, it is possible that white matter tracts connect the areas functionally, which are indirectly connected through the entorhinal cortex and parahippocampal cortex. A pathway between phonological areas and memory areas that does not pass through the lexical interface or articulatory network could perhaps help explain the modulatory effects of accent on memory performance.

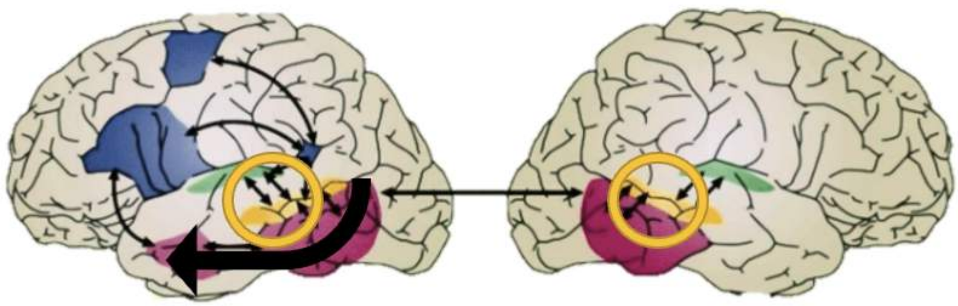


Fig. 28. Speculative area of accent modulation in the ventral stream of dual stream model. The phonological area is highlighted in the yellow circle. Image adapted from Hickok & Poeppel, 2007.

Taken as a whole, this doctoral work provides evidence that the human brain is highly sensitive to speech accent, even very early on in perception (chapter 3) and that there appears to be support that different types of accent affect different levels of cognitive processes differently. While foreign accent seems to uniquely impact early evoked responses (chapter 3) and high-frequency ranges (chapter 5), both dialectal accent and foreign accent appear to affect memory performance (chapter 4). Surprisingly, they had a beneficial effect on recognition memory performance, suggesting more needs to be explored especially in regards to how accents affect information retention over extended periods of time. While accent impacts an early stage of phonological analysis, it appears to be later resolved in the semantic processing phase. Despite this, this resolution appears to have lingering effects on memory. Based on the findings presented here, the study of speech accent processing would benefit from considering the possibility of flexible processing mechanisms that adapt along the time course of accent processing. Context-dependent models of accent processing highlight the adaptability of processing in response to non-native accented interlocutors. However, this adaptability may not be static, but rather a continuous process involving cognitive mechanisms that constantly update on a millisecond by-millisecond basis considering the social, linguistic, and cognitive context of the experience, as well as the real-time cognitive resources of the listener.

6.3. Limitations and future directions

Two of the most notable differences between the current dissertation and prior work examining accent processing were the methods of participant recruitment for one study and the multilingual area of the Basque Country. One experiment was adapted from its original in-lab design to an online experiment due to the COVID-19 pandemic. While online experiments are valuable and have been shown to provide reliable information (Luna et al., 2021), a concern with online data collection is the increased variability and decreased environmental control for the researcher than present in in-person data collection. While we attempted to limit variability by insisting upon the use of headphones and a private environment and implementing attentional reminders, we were limited in the amount we could control the conditions of the experiment in a home environment. When running a paradigm measuring attentional resource engagement, an ideal scenario would be to have a controlled environment where you could be sure not only that all conditions were the same but also that there were no additional unknown attentional burdens on the participants. Limiting online collection for this type of experiment and combining or replacing the dual-task paradigm with other promising attentional measures of listening effort, such as pupilometry, could serve as a valuable direction forward in understanding the amount of listening effort required for different accents. Which additionally may change in real-time with different accents, exposure levels, or participant profiles (cognitive resources, executive control, language background).

Another difference between the majority of the previous research in the field and the experiments conducted for this dissertation is that the Basque Center on Cognition, Brain, and Language is located in the Basque Region of Spain, which also crosses into France and thus is a multilingual area that is highly exposed to many languages and accents. While the participants were Spanish-dominant, it was likely that they are regularly exposed to a variety of languages and both dialectal and foreign accents. We did attempt to both limit familiarity to accents used in the study and collect information as covariates about exposure to dialectal and foreign accents in general but we were limited compared to popular areas of accent research such as the midwest of the United States (a very monolingual and often rural area). Indeed, previous work has shown that familiarity has an important effect on accent processing (Gass and Varonis 1984). In fact, listeners have been shown to adapt to accented speech especially if regularly exposed (Brown &

al., 2020; Evans & Taylor, 2010). The accent of a speaker can serve as a salient signal to the listener of whether their interlocutor is a member of their social group or not (see Campbell-Kibler, 2010; Drager, 2010). Stereotypes either consciously or even unconsciously associated with out-group members impact the credibility of the speaker and influence processing mechanisms in both beneficial and detrimental ways (Hosoda et al., 2007; Fuertes et al., 2002; Fiske, 1998). The influence on processing that foreign accent has is sometimes beneficial to preserve the fluidity of communication. Listeners adapt their processing of foreign-accented speech partially because of expectations associated with foreign-accented speech. While these assumptions can have negative effects such as activating prejudice or reduced credibility (Lev-Ari, 2017), they also serve a communicative function. This expectation-led adaptive processing allows listeners to be more flexible in their processing, more easily ignore errors, and increase their reliance on contextual information to aid comprehension (Lev-Ari & Peperkamp, 2014). However, it may also make differences between the processing of different accents more difficult to observe in areas especially exposed to accents. Despite this, the majority of the modern world is moving away from monolingualism in the traditional sense of the word and thus it is imperative that we continue to conduct research in culturally and linguistically diverse areas in order to understand how many people confront accent deviations in their naturally diverse linguistic environments.

6.4. Conclusion

In this chapter I have summarized the main findings of this doctoral work and how they contribute to the larger literature on language processing and accent specifically. In addition to its theoretical contribution, this work also has value in terms of practical applications. A complete understanding of the effects of indexical properties of the speech signal, and how they relate to cognitive processing differences, can be applied to the development of future more inclusive automatic speech recognition technologies (chapter 7) and strategies to mitigate (or take advantage of) effects due to speech accent variations. In recent years, research on accent processing has greatly advanced our understanding of the impact of accent on our perceptual system, but more work remains to be done to further understand the complexities of speech accent processing and lingering effects in diverse interactions. The evidence provided here suggests that the study of speech accent processing would benefit from considering the possibility of flexible processing mechanisms that adapt along the time course and cognitive level of accent processing. Future investigations may be better able to understand the intricacies of the flexibility of these processing mechanisms and the complex interplay between cognitive resources, adaptation, and social priming all associated with how we process non-native accents.

7. Related publications and manuscripts

In this section I will provide an overview of an ongoing project that I have developed in collaboration with Dr. Ying Xu at The University of California Irvine (UCI) and The University of Michigan during my research stay at UCI. Data analysis is still ongoing. However, preliminary results will be presented at the International Symposium on Bilingualism 2023. Publications and conference participation related to my dissertation are also listed in section 7.2.

7.1. Does bilingual status influence automatic speech recognition for young Latino children?

With the increased incorporation of Automatic Speech Recognition (ASR) systems in commonly used technologies, children are interacting with artificially intelligent systems from an early age. Millions of households around the world already have various smart technologies that employ ASR models in order to respond to information-seeking queries, order items, and manage other household needs. Recently, conversational agents using ASR have been shown to be both enjoyable for children to interact with (Druga et al., 2017)) and importantly, beneficial to learning outcomes (Xu & Warschauer, 2020; Xu et al., 2022). Xu et al., 2021 examined the use of technology using ASR during dialogic storybook reading and found that conversational agents using ASR can provide benefits comparable to a human partner. Critically, this finding supports that technology with ASR can provide an accessible alternative for children from families whose caretakers are not able to engage in dialogic co-reading as frequently to catch-up to their peers. However, this observed benefit was found only in the interactive condition and was not found when children were passively listening to the conversational agent during co-reading. Because of the importance of the verbal interaction between the conversational agent and the child, it is necessary to critically examine ASR performance with a broad and inclusive sample of speech in order to ensure that all children can benefit equally from current and future technologies.

It is well-documented that machine-learning technologies are susceptible to biases based on age, race, gender, etc. (Castro, 2019; Chouldechova, 2017; Obermeyer et al., 2019). One recent study looked at the transcription accuracy of several of the most-used ASR systems for speech-to-text transcriptions of the speech of White Americans and African Americans and found significant disparities in the word error rates of all ASR models when transcribing African

American speech (Koenecke et al., 2020). Additional analyses revealed that the disparities were most likely due to a performance gap in the acoustic models, suggesting that the phonological, phonetic or prosodic characteristics of African American Vernacular speech were confusing the models.

Bilingual children often must acquire two separate phonological systems simultaneously and although there is well-documented evidence for differentiated linguistic systems by age two, interaction, or transfer (the occurrence of sounds or sound patterns specific to one language in the other language context), between the two systems does occur (Barlow, 2003; Fabiano & Goldstein, 2005; Hossein Keshavarz & Ingram, 2002; Meisel, 1989). Many children speak a language other than English at home, the most common of which, in the United States, is Spanish. If previous studies have found evidence for a lack of inclusivity in the acoustic models of ASR, bilingual children, who are likely to have distinctive speech features, are at an increased risk of failing to benefit from educational advancements that rely on ASR. Furthermore, it is possible that disparities or differences in the way that bilingual children engage with ASR-integrated technology could be influenced by their environmental exposure to both languages. Performing a multifaceted investigation of the diverse sample of bilingual Latinos in the United States to understand how ASR performance is modulated by phonological features and the language background of bilingual children will help us to ensure that technologies are inclusively beneficial.

The principal objective of this research is to identify biases of ASR against Spanish-speaking bilingual children in the US and to explain the contributing factors of these biases, focusing on two broad questions:

1. What are the disparities in speech recognition performance i) between White, monolingual English-speaking children and Latino children who are exposed to English and Spanish; and ii) among Latino children with varying degrees of Spanish exposure and proficiency?

2. Are the observed disparities modulated by children's Spanish proficiency, linguistic environments, and acoustic features (i.e., Voice Onset Time)?

It has been shown that young language minority students, often Spanish heritage speakers, face additional educational challenges due to being inadequately prepared with early oral language skills (Irwin et al., 2021). Home-based interventions have been demonstrated as useful tools to mitigate these effects (Durán et al., 2015). This has led the Digital Learning Lab to investigate using conversational agents not only as an interactive learning partner but also as a tool to promote interaction between the child, the parent and the conversational agent

This project utilizes hundreds of speech samples from Spanish-English bilinguals with the aim of breaking down what specific mechanisms in the speech and linguistic environments affect the interactions of bilingual Latinos with ASR technologies. This project nicely complements ongoing projects of the Converse to Learn team of the Digital Learning Lab which is primarily concerned with developing educational tools to improve language development, literacy and STEM outcomes in low-income Latino children. Thus, this project can help inform ongoing and future projects of the Converse to Learn team's primary line of investigation which will affect millions of young digital technology users in the United States. Additionally, knowledge gained from this research can help diversify future ASR development and reduce biases that may hinder many people's ability to benefit from or enjoy new technological tools.

While previous research has identified racial disparities in ASR among adult speakers, our study will be the first rigorous study that directly focuses on ASR performance among Latino bilingual children, the fastest-growing ethnic group in the US. This study will provide valuable insight into the effectiveness of current ASR speech models for a diverse sample of children who are speakers of the second most spoken language in the United States and include students that are representative of the largest minority ethnicity in the United States.

With this project, we examined the effect of language proficiency (i.e., Bilingual English Spanish Oral Screener) of Spanish-English bilingual children from Southern California ($N = 200$; ages 3-7 years) on the transcription accuracy of their English speech by four widely used ASR systems (Amazon, Google, IBM, Microsoft). Speech samples were word-level speech collected via a picture-naming task in English. We computed the phonetic distance (Aline), and orthographic distance (Levenstein) between ASR transcription and human transcription. Preliminary analyses ($n = 137$) reveal differences in the accuracy of transcription between ASR systems. Critically, children who were older or more proficient in English had their speech transcribed more accurately by all APIs. Analyses on the larger sample will also examine how the quality of children's speech and cross-language influence from Spanish impact transcription using acoustic measures (i.e., Voice Onset Times). A group of age-matched, non-Hispanic, English monolingual children from the same region will provide a baseline for comparing transcription accuracy. Results highlight the need for greater language diversity in ASR model development.

Insights gained from this research can be used to help inform strategies for future technology development to reduce biases. Understanding how individual-level differences affect group-level benefits is critical to developing strategies for realizing inclusivity in ASR.

7.2. Related publications and conference participation

Journal Publications

Thomas, T., Martin, C.D., & Caffarra, S. (2022). An ERP investigation of accented isolated single word processing. Neuropsychologia, 175, 108349. doi:https://doi.org/10.1016/j.neuropsychologia.2022.108349.

Stoehr A., Souganidis, C., **Thomas, T**., Jacobsen, J., & Martin, C.D. (forthcoming). Voice onset time and vowel formant measures in online testing and laboratory-based testing with(out) surgical face masks. The Journal of the Acoustical Society of America.

Manuscripts in preparation

Xu, Y., **Thomas, T**., Takahesu-Tabori, A., Stoehr, A., Varady, C., & Warschaurer, M. (in preparation). Disparities in automated speech recognition among young bilingual children.

Thomas, T., Llorach, G., Martin, C., & Caffarra, S. (in preparation). Does accented speech affect attention and information retention?

Thomas, T., Martin, C., &Caffarra, S. (in preparation). A time-frequency investigation of accent processing in extended listening.

Stoehr A., Souganidis, C., **Thomas, T**., Jacobsen, J., & Martin, C.D. (forthcoming). Voice onset time and vowel formant measures in online testing and laboratory-based testing with(out) surgical face masks. The Journal of the Acoustical Society of America.

Stoehr A., Souganidis, C., **Thomas, T**., Jacobsen, J., & Martin, C.D. (in preparation). The speech elicitation method affects temporal but not spectral properties in trilinguals' speech production.

Peer-Reviewed Book Chapters

Thomas, T., Martin, C. D., Pesciarelli, F., Caffarra, S. (forthcoming). Analyzing Language using Brain Imaging. In P. Gygax and S. Zufferey (Eds.): The Routledge Handbook of Experimental Linguistics. Routledge.

Caffarra, S., Gosselin, L., **Thomas, T**., & Martin, C. (forthcoming). The Neurocognition of Foreign Accent Perception. Chapter 34, Part VII. The Routledge Handbook of Second Language Acquisition and Neurolinguistics.

Conference Proceeedings

Xu, Y., Levine, J., Vigil, V., Ritchie, D., **Thomas, T.**, Barrera, C., Meza, M., Zhang, S., Bustamante, A., & Warschauer, M. (2023, April). Conversations with media characters promote children's science learning from television watching [Paper presentation]. American Educational Research Association Annual Meeting, Chicago, IL.

Invited Talks

Stoehr A., Souganidis, C., **Thomas, T**., Jacobsen, J., & Martin, C.D. (2023, June). *Online testing and testing with facemasks are suitable for detecting language-specific VOT production in trilinguals*
Oral presentation at the International Symposium on Bilingualism (ISB14).

Thomas, T., Llorach, G., Martin, C., & Caffarra, S. (2022, September). *Does accented speech affect attention and information retention?*
Oral presentation at 4th International Symposium on Applied Phonetics, ISAPh 2022.

Thomas, T., (2022, March). *Speech recognition disparities among monolingual and bilingual Latine children.*
Invited speaker at the Brain Development and Education Lab meeting, Stanford University.

Thomas, T. (2021, December). *Exploring the effects of accent on cognitive processes.*
Public presentation to University of California Irvine affiliates, invited by Digital Learning Lab of the university.

Poster Presentations

Thomas, T., Bernardi, I., Martin, C.D., & Caffarra, S. (2023, August). *Online Processing Of Native, Dialectal And Foreign Accent During Extended Listening Tasks.*
ESCoP conference, 23rd Conference of the European Society for Cognitive Psychology.

Thomas, T., Bernardi, I., Martin, C.D., & Caffarra, S. (2023, August). *A time-frequency investigation of native, dialectal and foreign accent processing.*
29th AMLaP conference, Architectures and Mechanisms for Language Processing.

Thomas, T., Takahesu-Tabori, A., Stoehr, A., Varady, C., & Xu, Y. (2023, June). *Does bilingual status influence automatic speech recognition for young Latino children?*
International Symposium on Bilingualism (ISB14).

Stoehr A., Souganidis, C., **Thomas, T**., Jacobsen, J., & Martin, C.D. (2023, May). Phonetic detail in trilinguals' speech production in online testing and testing with face masks. Internatoinal Symposium of Psycholinguistics.

Thomas, T., Llorach, G., Martin, C., & Caffarra, S. (2022, August). *Listening effort across accents: Does it affect memory?*
22nd Conference of the European Society for Cognitive Psychology (ESCoP) 2022.

Thomas, T., Martin, C., & Caffarra, S. (2021, March). *Towards a dynamic model of processing of native, dialectal, and foreign accented speech*.
Cognitive Neuroscience Society (CNS) 2021Virtual.

8. References

Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analyses. *Brain and Language*, *122*(1), 42–54. https://doi.org/10.1016/j.bandl.2012.04.014

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(2), 520–529. https://doi.org/10.1037/a0013552

Adank, P., & Janse, E. (2009). Perceptual learning of time-compressed and natural fast speech. *The Journal of the Acoustical Society of America*, *126*(5), 2649–2659. https://doi.org/10.1121/1.3216914

Adank, P., & McQueen, J. M. (2007). The effect of an unfamiliar regional accent on spoken-word comprehension. *16th International Congress of Phonetic Sciences (ICPhS 2007)*, 1925–1928. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_59662

Alshemali, B., & Kalita, J. (2020). Improving the Reliability of Deep Neural Networks in NLP: A Review. *Knowledge-Based Systems*, *191*, 105210. https://doi.org/10.1016/j.knosys.2019.105210

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, *38*(4), 561–613. https://doi.org/10.1111/j.1467-1770.1988.tb00167.x

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent’s voice. *Contemporary Educational*

Psychology, *30*(1), 117–139. https://doi.org/10.1016/j.cedpsych.2004.07.001

Baese-Berk, M. M., McLaughlin, D. J., & McGowan, K. B. (2020). Perception of non‐native speech. *Language and Linguistics Compass*, *14*(7). https://doi.org/10.1111/lnc3.12375

Banaji, M. R., & Hardin, C. D. (1996). Automatic Stereotyping. *Psychological Science*, *7*(3), 136–141. https://doi.org/10.1111/j.1467-9280.1996.tb00346.x

Barlow, J. A. (2003). The stop-spirant alternation in Spanish: converging evidence for a fortition account. *Southwest Journal of Linguistics*, *22*, 51+. https://go.gale.com/ps/i.do?id=GALE%7CA109268631&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=07374143&p=AONE&sw=w

Bastiaansen, M. C. M., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language*, *106*(1), 15–28. https://doi.org/10.1016/j.bandl.2007.10.006

Bastiaansen, M. C. M., Posthuma, D., Groot, P. F. C., & de Geus, E. J. C. (2002). Event-related alpha and theta responses in a visuo-spatial working memory task. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *113*(12), 1882–1893. https://doi.org/10.1016/s1388-2457(02)00303-6

Baus, C., Bas, J., Calabria, M., & Costa, A. (2017). Speak My Language and I Will Remember Your Face Better: An ERP Study. *Frontiers in Psychology*, *8*, 709. https://doi.org/10.3389/fpsyg.2017.00709

Bentin, S., Kutas, M., & Hillyard, S. A. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology*, *30*(2), 161–169. https://doi.org/10.1111/j.1469-8986.1993.tb01729.x

Bent, T., Bradlow, A. R., & Wright, B. A. (2003). The influence of linguistic experience on pitch perception in speech and nonspeech sounds. *The Journal of the Acoustical Society of America*, *113*(4), 2256–2256. https://doi.org/10.1121/1.4780440

Berger, H. (1929). Über das elektroenkephalogramm des menschen. *Archiv Für Psychiatrie Und Nervenkrankheiten*, *87*(1), 527–570. https://pure.mpg.de/rest/items/item_2281721/component/file_2281720/content

Best, C. T., & Others. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, *167*(224), 233–277. https://www.academia.edu/download/45189090/The_emergence_of_native-language_phonolo20160428-29981-atkgy1.pdf

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, *56*, 9–49. https://doi.org/10.1111/j.1467-9922.2006.00353.x

Borghini, G., & Hazan, V. (2018). Listening Effort During Sentence Processing Is Increased for Non-native Listeners: A Pupillometry Study. *Frontiers in Neuroscience*, *12*, 152. https://doi.org/10.3389/fnins.2018.00152

Bourguignon, M., De Tiège, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert, P., Goldman, S., Hari, R., & Jousmäki, V. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human Brain Mapping*, *34*(2), 314–326. https://doi.org/10.1002/hbm.21442

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. 'ichi. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977–985. https://doi.org/10.3758/BF03206911

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299–2310. https://doi.org/10.1121/1.418276

Braun, B., Lemhöfer, K., & Mani, N. (2011). Perceiving unstressed vowels in foreign-accented English. *The Journal of the Acoustical Society of America*, *129*(1), 376–387. https://doi.org/10.1121/1.3500688

Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin et Memoires de La Societe Anatomique de Paris*, *6*, 330–357. https://pure.mpg.de/rest/items/item_2301213/component/file_2452907/content

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience*, *7*, 758. https://doi.org/10.3389/fnhum.2013.00758

Brown, V. A., McLaughlin, D. J., Strand, J. F., & Van Engen, K. J. (2020a). Rapid adaptation to fully intelligible nonnative-accented speech reduces listening effort. *Quarterly Journal of Experimental Psychology* , *73*(9), 1431–1443. https://doi.org/10.1177/1747021820916726

Brown, V. A., McLaughlin, D. J., Strand, J. F., & Van Engen, K. J. (2020b). Intelligible nonnative-accented speech incurs a processing cost: Converging evidence from a behavioral and a physiological paradigm. *The Journal of the Acoustical Society of America*, *148*(4), 2653–2653. https://doi.org/10.1121/1.5147383

Bürki-Cohen, J., Miller, J. L., & Eimas, P. D. (2001). Perceiving non-native speech. *Language and Speech*, *44*(Pt 2), 149–169. https://doi.org/10.1177/00238309010440020201

Cacciari, C., & Padovani, R. (2007). Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics*, *28*(2), 277–293. https://doi.org/10.1017/S0142716407070142

Caffarra, S., & Martin, C. D. (2019). Not all errors are the same: ERP sensitivity to error typicality in foreign accented speech perception. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *116*, 308–320. https://doi.org/10.1016/j.cortex.2018.03.007

Casserly, E. D., & Pisoni, D. B. (2010). Speech perception and production. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(5), 629–647. https://doi.org/10.1002/wcs.63

Castro, C. (2019). What's wrong with machine bias. *Ergo, an Open Access Journal of Philosophy*, *6*. https://quod.lib.umich.edu/e/ergo/12405314.0006.015/--what-s-wrong-with-machine-bias?rgn=main;view=fulltext

Cheng, L. S. P., Burgess, D., Vernooij, N., Solís-Barroso, C., McDermott, A., & Namboodiripad, S. (2021). The Problematic Concept of Native Speaker in Psycholinguistics: Replacing Vague and Harmful Terminology With Inclusive and Accurate Measures. *Frontiers in Psychology*, *12*, 715843. https://doi.org/10.3389/fpsyg.2021.715843

Chen, J., & Chang, H. (2022). Sketching the Landscape of Speech Perception Research (2000-2020): A Bibliometric Study. *Frontiers in Psychology*, *13*, 822241. https://doi.org/10.3389/fpsyg.2022.822241

Cho, K. W., & Feldman, L. B. (2013). Production and accent affect memory. *The Mental Lexicon*, *8*(3), 295–319. https://doi.org/10.1075/ml.8.3.02cho

Chouldechova, A. (2017). Fair Prediction with Disparate Impact: A Study of Bias in Recidivism

Prediction Instruments. *Big Data*, *5*(2), 153–163. https://doi.org/10.1089/big.2016.0047

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647–3658. https://doi.org/10.1121/1.1815131

Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: intelligibility and classification. *Language and Speech*, *51*(Pt 3), 175–198. https://doi.org/10.1177/0023830908098539

Cohen, J. (1977). The Concepts of Power Analysis. In *Statistical Power Analysis for the Behavioral Sciences* (pp. 1–17). https://doi.org/10.1016/b978-0-12-179060-8.50006-2

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, *6*(3), 256–266. https://doi.org/10.1162/jocn.1994.6.3.256

Connolly, J. F., Service, E., D'Arcy, R. C., Kujala, A., & Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *Neuroreport*, *12*(2), 237–243. https://doi.org/10.1097/00001756-200102120-00012

Credibility of native and non-native speakers of English revisited: Do non-native listeners feel the same? (2017). *Research in Language (RiL)*, *15*(3), 285–298. https://www.ceeol.com/search/article-detail?id=750375

Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, *3*, 479. https://doi.org/10.3389/fpsyg.2012.00479

Dave, S., Brothers, T., Hoversten, L. J., Traxler, M. J., & Swaab, T. Y. (2021). Cognitive control mediates age-related changes in flexible anticipatory processing during listening comprehension. *Brain Research*, *1768*, 147573. https://doi.org/10.1016/j.brainres.2021.147573

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, *38*(2), 105–134. https://doi.org/10.1007/s11251-009-9110-0

Desjardins, J. L., & Doherty, K. A. (2014). The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear and Hearing*, *35*(6), 600–610. https://doi.org/10.1097/AUD.0000000000000028

Dickter, C. L., Kittel, J. A., & Gyurovski, I. I. (2012). Perceptions of non-target confronters in response to racist and heterosexist remarks. *European Journal of Social Psychology*, *42*(1), 112–119. https://doi.org/10.1002/ejsp.855

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85 Pt 2*(0 2), 761–768. https://doi.org/10.1016/j.neuroimage.2013.06.035

Druga, S., Williams, R., Breazeal, C., & Resnick, M. (2017, June 27). Hey Google is it OK if I eat you? *Proceedings of the 2017 Conference on Interaction Design and Children*. IDC '17: Interaction Design and Children, Stanford California USA. https://doi.org/10.1145/3078072.3084330

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: one-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Durán, R., Estay-Niculcar, C., & Álvarez, H. (2015). Adopción de buenas prácticas en la educación virtual en la educación superior. *Aula Abierta*, *43*(2), 77–86.

https://doi.org/10.1016/j.aula.2015.01.001

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. In *Behavior Research Methods, Instruments, & Computers* (Vol. 28, Issue 1, pp. 1–11). https://doi.org/10.3758/bf03203630

Evans, B. G., & Taylor, E. C. (n.d.). Investigating the time course of perceptual adaptation to unfamiliar accented speech. *Proceedings of BAAP 2010*.

Fabiano, L., & Goldstein, B. (2005). Phonological cross-linguistic effects in bilingual Spanish – English speaking children. *Journal of Multilingual Communication Disorders*, *3*(1), 56–63. https://doi.org/10.1080/14769670400027316

Fitch, R. H., Miller, S., & Tallal, P. (1997). Neurobiology of speech perception. *Annual Review of Neuroscience*, *20*, 331–353. https://doi.org/10.1146/annurev.neuro.20.1.331

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, *92*, 233–277. https://www.researchgate.net/profile/James-Flege/publication/333815781_Second_language_speech_learning_Theory_findings_and_problems/links/5d071d2692851c900442d6b2/Second-language-speech-learning-Theory-findings-and-problems.pdf

Floccia, C., Butler, J., Girard, F., & Goslin, J. (2009). Categorization of regional and foreign accent in 5- to 7-year-old British children. *International Journal of Behavioral Development*, *33*(4), 366–375. https://doi.org/10.1177/0165025409103871

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology. Human Perception and Performance*, *32*(5), 1276–1293. https://doi.org/10.1037/0096-1523.32.5.1276

Foucart, A., Costa, A., Morís-Fernández, L., & Hartsuiker, R. J. (2020). Foreignness or processing fluency? On understanding the negative bias toward foreign‐accented speakers. *Language Learning*, *70*(4), 974–1016. https://doi.org/10.1111/lang.12413

Foucart, A., & Hartsuiker, R. J. (2021). Are foreign-accented speakers that "incredible"? The impact of the speaker's indexical properties on sentence processing. *Neuropsychologia*, *158*, 107902. https://doi.org/10.1016/j.neuropsychologia.2021.107902

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). How do you know I was about to say "book"? Anticipation processes affect speech processing and lexical recognition. *Language, Cognition and Neuroscience*, *30*(6), 768–780. https://doi.org/10.1080/23273798.2015.1016047

Fougnie, D. (n.d.). *The relationship between attention and working memory*. Retrieved February 17, 2023, from https://blogs.lt.vt.edu/devineramblings/files/2018/11/Attention-and-Memory.pdf

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*(4), 409–438. https://doi.org/10.1016/j.wocn.2005.08.002

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct–realist perspective. *Journal of Phonetics*, *14*(1), 3–28. https://doi.org/10.1016/S0095-4470(19)30607-2

Frances, C., Costa, A., & Baus, C. (2018). On the effects of regional accents on memory and credibility. *Acta Psychologica*, *186*, 63–70. https://doi.org/10.1016/j.actpsy.2018.04.003

Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, *17*(3), 213–247. https://doi.org/10.1177/026765830101700301

Franceschina, F. (2005). *Fossilized Second Language Grammars: The Acquisition of*

Grammatical Gender. John Benjamins Publishing. https://play.google.com/store/books/details?id=54H1V6SU1C0C

Gagné, J.-P., Besser, J., & Lemke, U. (2017). Behavioral Assessment of Listening Effort Using a Dual-Task Paradigm. *Trends in Hearing*, *21*, 2331216516687287. https://doi.org/10.1177/2331216516687287

Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*(1), 65–87. https://doi.org/10.1111/j.1467-1770.1984.tb00996.x

Girard, F., Floccia, C., & Goslin, J. (2008). Perception and awareness of accents in young children. *The British Journal of Developmental Psychology*, *26*(3), 409–433. https://doi.org/10.1348/026151007x251712

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(5), 1166–1183. https://doi.org/10.1037//0278-7393.22.5.1166

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279. https://doi.org/10.1037/0033-295x.105.2.251

Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language*, *122*(2), 92–102. https://doi.org/10.1016/j.bandl.2012.04.017

Gosselin, L., Martin, C. D., Navarra-Barindelli, E., & Caffarra, S. (2021). The presence of a foreign accent introduces lexical integration difficulties during late semantic processing. *Language, Cognition and Neuroscience*, *36*(9), 1086–1106. https://doi.org/10.1080/23273798.2021.1909084

Gow, D. W., Jr. (2012). The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. *Brain and Language*, *121*(3), 273–288. https://doi.org/10.1016/j.bandl.2012.03.005

Grabot, L., Kösem, A., Azizi, L., & van Wassenhove, V. (2017). Prestimulus Alpha Oscillations and the Temporal Sequencing of Audiovisual Events. In *Journal of Cognitive Neuroscience* (Vol. 29, Issue 9, pp. 1566–1582). https://doi.org/10.1162/jocn_a_01145

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/bf02289823

Grey, S., Cosgrove, A. L., & van Hell, J. G. (2020). Faces with foreign accents: An event-related potential study of accented sentence comprehension. *Neuropsychologia*, *147*, 107575. https://doi.org/10.1016/j.neuropsychologia.2020.107575

Grey, S., Schubel, L. C., Mcqueen, J. M., & Van Hell, J. G. (2019). Processing foreign-accented speech in a second language: Evidence from ERPs during sentence comprehension in bilinguals. *Bilingualism: Language and Cognition*, *22*(5), 912–929. https://doi.org/10.1017/S1366728918000937

Grey, S., & van Hell, J. G. (2017). Foreign-accented speaker identity affects neural correlates of language comprehension. *Journal of Neurolinguistics*, *42*, 93–108. https://doi.org/10.1016/j.jneuroling.2016.12.001

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology*, *11*(12), e1001752. https://doi.org/10.1371/journal.pbio.1001752

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, *21*(2), 205–228. https://www.cambridge.org/core/journals/applied-psycholinguistics/article/age-of-learning-effects-on-the-duration-of-sentences-produced-in-a-second-language/785BF7DBC4B824B87B7593E9E2E9F612

Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 801–811. https://doi.org/10.1098/rstb.2007.2089

Hämäläinen, M., & Hari, R. (2002). Magnetoencephalographic (MEG) characterization of dynamic brain activation. *Brain Mapping: The Methods, Ed*, *2*, 227–254. http://kutaslab.ucsd.edu/people/kutas/279/pdfs/2004.BMTM.1.pdf

Hansen, K., Steffens, M. C., Rakic, T., & Wiese, H. (2017). When appearance does not match accent: neural correlates of ethnicity-related expectancy violations. *Social Cognitive and Affective Neuroscience*, *12*(3), 507–515. https://doi.org/10.1093/scan/nsw148

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person’s mistake is another's standard usage: the effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, *24*(4), 878–887. https://doi.org/10.1162/jocn_a_00103

Hatzidaki, A., Baus, C., & Costa, A. (2015). The way you say it, the way I feel it: emotional word processing in accented speech. *Frontiers in Psychology*, *6*, 351. https://doi.org/10.3389/fpsyg.2015.00351

Hay, J., Nolan, A., & Drager, K. (2006). *From fush to feesh: Exemplar priming in speech perception*. *23*(3), 351–379. https://doi.org/10.1515/TLR.2006.014

Heinrich, A., Schneider, B. A., & Craik, F. I. M. (2008). Investigating the influence of continuous babble on auditory short-term memory performance. *Quarterly Journal of Experimental Psychology* , *61*(5), 735–751. https://doi.org/10.1080/17470210701402372

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

Hickok, G., & Poeppel, D. (2016). Chapter 25 - Neural Basis of Speech Perception. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 299–310). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00025-0

Hossein Keshavarz, M., & Ingram, D. (2002). The early phonological development of a Farsi-English bilingual child. *International Journal of Bilingualism*, *6*(3), 255–269. https://doi.org/10.1177/13670069020060030301

Howard, M. F., & Poeppel, D. (2012). The neuromagnetic response to spoken sentences: co-modulation of theta band amplitude and phase. *NeuroImage*, *60*(4), 2118–2127. https://doi.org/10.1016/j.neuroimage.2012.02.028

Irwin, V., Zhang, J., Wang, X., Hein, S., Wang, K., Roberts, A., York, C., Barmer, A., Bullock Mann, F., Dilig, R., & Others. (2021). Report on the Condition of Education 2021. NCES 2021-144. *National Center for Education Statistics*. https://files.eric.ed.gov/fulltext/ED612942.pdf

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the Alpha Band (9–12 Hz) Increase with Memory Load during Retention in a Short-term Memory Task. *Cerebral Cortex* , *12*(8), 877–882. https://doi.org/10.1093/cercor/12.8.877

Jiang, X., Gossack-Keenan, K., & Pell, M. D. (2020). To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Quarterly Journal of Experimental Psychology* , *73*(1), 55–79. https://doi.org/10.1177/1747021819865833

Johnson, K. (1997). *The auditory/perceptual basis for speech segmentation*. *101*, 113. https://kb.osu.edu/handle/1811/81782

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*(4), 359–384. https://doi.org/10.1006/jpho.1999.0100

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(3), 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Kavas, A., & Kavas, A. (2008). An exploratory study of undergraduate college students' perceptions and attitudes toward foreign accented faculty. *College Student Journal*, *42*, 879+. https://go.gale.com/ps/i.do?id=GALE%7CA182975282&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=01463934&p=AONE&sw=w

Kayte, S., Mundada, M., & Gujrathi, J. (2015). Hidden Markov Model based Speech Synthesis: A Review. *International Journal of Computers & Applications*, *130*(3), 35–39. https://doi.org/10.5120/ijca2015906965

Kemper, S., Schmalzried, R., Herman, R., Leedahl, S., & Mohankumar, D. (2009). The effects of aging and dual task demands on language production. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *16*(3), 241–259. https://doi.org/10.1080/13825580802438868

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(14), 7684–7689. https://doi.org/10.1073/pnas.1915768117

Krause, C. M., Lang, A. H., Laine, M., Kuusisto, M., & Pörn, B. (1996). Event-related EEG desynchronization and synchronization during an auditory memory task. *Electroencephalography and Clinical Neurophysiology*, *98*(4), 319–326. https://doi.org/10.1016/0013-4694(96)00283-0

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

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Kutas, M., Neville, H. J., & Holcomb, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening and signing. *Electroencephalography and Clinical Neurophysiology. Supplement*, *39*, 325–330.

https://www.ncbi.nlm.nih.gov/pubmed/3477442

Labov, W., Ash, S., & Boberg, C. (n.d.). A national map of the regional dialects of American English. *Linguistics Laboratory, Dept. of Linguistics, U of.*

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, *60*, 44–51. https://doi.org/10.1037/h0044430

Lane, H. (1963). Foreign Accent and Speech Distortion. *The Journal of the Acoustical Society of America*, *35*(4), 451–453. https://doi.org/10.1121/1.1918501

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews. Neuroscience*, *9*(12), 920–933. https://doi.org/10.1038/nrn2532

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A.-L. (2011). Altered low-γ sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron*, *72*(6), 1080–1090. https://doi.org/10.1016/j.neuron.2011.11.002

Leiberg, S., Lutzenberger, W., & Kaiser, J. (2006). Effects of memory load on cortical oscillatory activity during auditory pattern working memory. *Brain Research*, *1120*(1), 131–140. https://doi.org/10.1016/j.brainres.2006.08.066

Lev-Ari, S., & Keysar, B. (2012). Less-Detailed Representation of Non-Native Language: Why Non-Native Speakers' Stories Seem More Vague. *Discourse Processes*, *49*(7), 523–538. https://doi.org/10.1080/0163853X.2012.698493

Lev-Ari, S., & Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception☆. *Journal of Phonetics*, *47*, 36–46. https://doi.org/10.1016/j.wocn.2014.09.001

Lewis, A. G., & Bastiaansen, M. (2015). *A predictive coding framework for rapid neural* *dynamics during sentence-level language comprehension. Cortex, 68, 155e168*.

Lewis, A. G., Schoffelen, J.-M., Schriefers, H., & Bastiaansen, M. (2016). A Predictive Coding Perspective on Beta Oscillations during Sentence-Level Language Comprehension. *Frontiers in Human Neuroscience*, *10*, 85. https://doi.org/10.3389/fnhum.2016.00085

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1–36. https://doi.org/10.1016/0010-0277(85)90021-6

Luce, P. A., & Mclennan, C. T. (n.d.). *SSpoken word recognition: The challenge of VSSpoken word recognition: The challenge of variation ariation*. Retrieved June 5, 2023, from https://engagedscholarship.csuohio.edu/cgi/viewcontent.cgi?article=1016&context=clpsych_facpub

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique, second edition*. MIT Press. https://play.google.com/store/books/details?id=SzavAwAAQBAJ

Luna, F. G., Roca, J., Martín-Arévalo, E., & Lupiáñez, J. (2021). Measuring attention and vigilance in the laboratory vs. online: The split-half reliability of the ANTI-Vea. *Behavior Research Methods*, *53*(3), 1124–1147. https://doi.org/10.3758/s13428-020-01483-4

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Mai, G., Minett, J. W., & Wang, W. S.-Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, *133*, 516–528. https://doi.org/10.1016/j.neuroimage.2016.02.064

Major, B., Quinton, W. J., & McCoy, S. K. (2002). Antecedents and consequences of attributions to discrimination: Theoretical and empirical advances. In *Advances in Experimental Social*

Psychology (Vol. 34, pp. 251–330). Academic Press. https://doi.org/10.1016/S0065-2601(02)80007-7

Maltseva, I., Geissler, H. G., & Başar, E. (2000). Alpha oscillations as an indicator of dynamic memory operations - anticipation of omitted stimuli. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *36*(3), 185–197. https://doi.org/10.1016/s0167-8760(99)00093-8

Marden, C. C. (1896). IV.—The Phonology of the Spanish Dialect of Mexico City. *PMLA*, *11*(1), 85–150. https://doi.org/10.2307/456218

Mariko, A. (2007). Grammatical errors across proficiency levels in L2 spoken and written English. *The Economic Journal of Takasaki City University of Economics*, *49*(3), 4. https://www.academia.edu/download/34614472/Abe2007_SPWR.pdf

Mason, H. M. (1946). Understandability of speech in noise as affected by region of origin of speaker and listener. *Speech Monographs*, *13*(2), 54–58. https://doi.org/10.1080/03637754609374918

Mateu, R. O. (1996). Comparación fonética, diagnóstico y tratamiento de las dificultades de los estudiantes chinos para aprender español. *Aspectos de didáctica de la lengua y la literatura: actas del III Congreso Internacional de la Sociedad española de didáctica de la lengua y la literatura*, 319–334. https://dialnet.unirioja.es/servlet/articulo?codigo=1037571

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology. Applied*, *11*(4), 256–265. https://doi.org/10.1037/1076-898X.11.4.256

Mayer, R. E., Sobko, K., & Mautone, P. D. (n.d.). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology*, *95*(2), 419–425. https://doi.org/10.1037/0022-0663.95.2.419

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say "hello"? Personality impressions from brief novel voices. *PloS One*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

McLaughlin, D. J., Braver, T. S., & Peelle, J. E. (2021). Measuring the Subjective Cost of Listening Effort Using a Discounting Task. *Journal of Speech, Language, and Hearing Research: JSLHR*, *64*(2), 337–347. https://doi.org/10.1044/2020_JSLHR-20-00086

McLaughlin, D. J., & Van Engen, K. J. (2020). Task-evoked pupil response for accurately recognized accented speech. *The Journal of the Acoustical Society of America*, *147*(2), EL151. https://doi.org/10.1121/10.0000718

Meisel, J. M. (1989). *Early differentiation of languages in bilingual children. In. K. Hyltenstam e L. Obler (orgs.) Bilingualism Across the Life Span: Aspects of Acquisition, Maturity, and Loss*. Cambridge: Cambridge University Press.

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2017). Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cerebral Cortex* , *27*(9), 4293–4302. https://doi.org/10.1093/cercor/bhw228

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., de Lissa, P., Graham, P., & Lyxell, B. (2017). Objective Assessment of Listening Effort: Coregistration of Pupillometry and EEG. *Trends in Hearing*, *21*, 2331216517706396. https://doi.org/10.1177/2331216517706396

Mok, P., & Dellwo, V. (2008). *Comparing native and non-native speech rhythm using acoustic* *rhythmic measures: Cantonese, Beijing Mandarin and English*. 423–426. https://doi.org/10.5167/uzh-111798

Munro, M. J., & Derwing, T. M. (1995a). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38 (Pt 3)*, 289–306. https://doi.org/10.1177/002383099503800305

Munro, M. J., & Derwing, T. M. (1995b). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Newman, R. L., & Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom–up and top–down effects on spoken word processing. *Biological Psychology*, *80*(1), 114–121. https://doi.org/10.1016/j.biopsycho.2008.04.008

Newman, R. L., Connolly, J. F., Service, E., & McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: evidence from event-related brain potentials. *Psychophysiology*, *40*(4), 640–647. https://doi.org/10.1111/1469-8986.00065

Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2004). Stimulus modality, perceptual overlap, and the go/no-go N2. *Psychophysiology*, *41*(1), 157–160. https://doi.org/10.1046/j.1469-8986.2003.00128.x

Nikolov, M., & Djigunović, J. M. (2006). RECENT RESEARCH ON AGE, SECOND LANGUAGE ACQUISITION, AND EARLY FOREIGN LANGUAGE LEARNING. *Annual Review of Applied Linguistics*, *26*, 234–260. https://doi.org/10.1017/S0267190506000122

Nygaard, L. C., & Pisoni, D. B. (1995). Speech perception: New directions in research and theory. *Speech, Language, and Communication.*, *415*, 63–96.

https://doi.org/10.1016/B978-012497770-9/50005-4

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, *60*(3), 355–376. https://doi.org/10.3758/bf03206860

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). SPEECH PERCEPTION AS A TALKER-CONTINGENT PROCESS. *Psychological Science*, *5*(1), 42–46. https://doi.org/10.1111/j.1467-9280.1994.tb00612.x

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Oosthuizen, I., Picou, E. M., Pottas, L., Myburgh, H. C., & Swanepoel, D. W. (2020). Listening Effort in Native and Nonnative English-Speaking Children Using Low Linguistic Single- and Dual-Task Paradigms. *Journal of Speech, Language, and Hearing Research: JSLHR*, *63*(6), 1979–1989. https://doi.org/10.1044/2020_JSLHR-19-00330

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237–241. https://doi.org/10.1016/j.tics.2008.02.014

Ortiz-Mantilla, S., Hämäläinen, J. A., Musacchia, G., & Benasich, A. A. (2013). Enhancement of gamma oscillations indicates preferential processing of native over foreign phonemic contrasts in infants. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(48), 18746–18754. https://doi.org/10.1523/JNEUROSCI.3260-13.2013

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. In *Educational Psychologist* (Vol. 38, Issue 1, pp. 1–4). https://doi.org/10.1207/s15326985ep3801_1

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive Load

Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2016). Cognitive load measurement as a means to advance cognitive load theory. In *Educational psychologist* (pp. 63–71). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780203764770-8/cognitive-load-measurement-means-advance-cognitive-load-theory-fred-paas-juhani-tuovinen-huib-tabbers-pascal-van-gerven

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex* , *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Peña, M., & Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of Cognitive Neuroscience*, *24*(5), 1149–1164. https://doi.org/10.1162/jocn_a_00144

Peng, Z. E., & Wang, L. M. (2019). Listening Effort by Native and Nonnative Listeners Due to Noise, Reverberation, and Talker Foreign Accent During English Speech Perception. *Journal of Speech, Language, and Hearing Research: JSLHR*, *62*(4), 1068–1081. https://doi.org/10.1044/2018_JSLHR-H-17-0423

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, *37 Suppl 1*, 5S – 27S. https://doi.org/10.1097/AUD.0000000000000312

Pickel, K. L., & Staller, J. B. (2012). A perpetrator's accent impairs witnesses' memory for physical appearance. *Law and Human Behavior*, *36*(2), 140–150. https://doi.org/10.1037/h0093968

Porretta, V., Tremblay, A., & Bolger, P. (2017). Got experience? PMN amplitudes to foreign-accented speech modulated by listener experience. *Journal of Neurolinguistics*, *44*, 54–67. https://doi.org/10.1016/j.jneuroling.2017.03.002

Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *The Quarterly Journal of Experimental Psychology*, *20*(3), 241–248. https://doi.org/10.1080/14640746808400158

Ranganath, C., Heller, A., Cohen, M. X., Brozinsky, C. J., & Rissman, J. (2005). Functional connectivity with the hippocampus during successful memory formation. *Hippocampus*, *15*(8), 997–1005. https://doi.org/10.1002/hipo.20141

Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Brain Research. Cognitive Brain Research*, *17*(3), 781–791. https://doi.org/10.1016/s0926-6410(03)00202-7

Rogers, C. L., Dalby, J., & Nishi, K. (2004). Effects of noise and proficiency on intelligibility of Chinese-accented English. *Language and Speech*, *47*(Pt 2), 139–154. https://doi.org/10.1177/00238309040470020201

Romero-Rivas, C., Martin, C. D., & Costa, A. (2015). Processing changes when listening to foreign-accented speech. *Frontiers in Human Neuroscience*, *9*, 167. https://doi.org/10.3389/fnhum.2015.00167

Romero-Rivas, C., Martin, C. D., & Costa, A. (2016). Foreign-accented speech modulates linguistic anticipatory processes. *Neuropsychologia*, *85*, 245–255.

https://doi.org/10.1016/j.neuropsychologia.2016.03.022

Romero-Rivas, C., Morgan, C., & Collier, T. (2022). Accentism on Trial: Categorization/Stereotyping and Implicit Biases Predict Harsher Sentences for Foreign-Accented Defendants. *Journal of Language and Social Psychology*, *41*(2), 191–208. https://doi.org/10.1177/0261927X211022785

Rossi, S., Gugler, M., Friederici, A., & Hahne, A. (n.d.). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. Cognitive …. *Journal of Atomic and Molecular Physics*.

Ryan, E. B., Giles, H., & Sebastian, R. J. (n.d.). An integrative perspective for the study of attitudes toward language variation. *Variation: Social and …*.

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, *37*(2), 217–240. https://doi.org/10.1017/S0142716414000502

Samuel, A. G. (2011). Speech perception. *Annual Review of Psychology*, *62*, 49–72. https://doi.org/10.1146/annurev.psych.121208.131643

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., Gruber, W. R., & Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *The European Journal of Neuroscience*, *22*(11), 2917–2926. https://doi.org/10.1111/j.1460-9568.2005.04482.x

Schiller, N. O., Boutonnet, B. P.-A., De Heer Kloots, M. L. S., Meelen, M., Ruijgrok, B., & Cheng, L. L.-S. (2020). (Not so) Great Expectations: Listening to Foreign-Accented Speech Reduces the Brain's Anticipatory Processes. *Frontiers in Psychology*, *11*, 2143.

https://doi.org/10.3389/fpsyg.2020.02143

Schmid, P. M., & Yeni-Komshian, G. H. (1999). The effects of speaker accent and target predictability on perception of mispronunciations. *Journal of Speech, Language, and Hearing Research: JSLHR*, *42*(1), 56–64. https://doi.org/10.1044/jslhr.4201.56

Schmidt-Kassow, M., & Kotz, S. A. (2008). Entrainment of syntactic processing? ERP-responses to predictable time intervals during syntactic reanalysis. *Brain Research*, *1226*, 144–155. https://doi.org/10.1016/j.brainres.2008.06.017

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law: An Official Law Review of the University of Arizona College of Law and the University of Miami School of Law*. https://psycnet.apa.org/journals/law/7/1/36/

Stevenage, S. V., Clarke, G., & McNeill, A. (2012). The effect of regional accent on voice recognition. *Journal of Cognitive Psychology* . https://eprints.soton.ac.uk/336628/

Strand, J. F., Brown, V. A., Merchant, M. B., Brown, H. E., & Smith, J. (2018). Measuring Listening Effort: Convergent Validity, Sensitivity, and Links With Cognitive and Personality Measures. *Journal of Speech, Language, and Hearing Research: JSLHR*, *61*(6), 1463–1486. https://doi.org/10.1044/2018_JSLHR-H-17-0257

Sumner, M., & Tilsen, S. (2011). Low-level adjustments to gross-categorical mismatches in the perception of accented speech. *The Journal of the Acoustical Society of America*, *130*(4), 2571–2571. https://doi.org/10.1121/1.3655300

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, *18*(1), 70–73. https://doi.org/10.4103/0972-6748.57865

Thomas, T., Martin, C. D., & Caffarra, S. (2022). An ERP investigation of accented isolated single word processing. *Neuropsychologia*, *175*, 108349. https://doi.org/10.1016/j.neuropsychologia.2022.108349

Tun, P. A., Wingfield, A., & Stine, E. A. (1991). Speech-processing capacity in young and older adults: a dual-task study. *Psychology and Aging*, *6*(1), 3–9. https://doi.org/10.1037//0882-7974.6.1.3

Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology. Human Learning and Memory*, *5*(6), 607. https://psycnet.apa.org/journals/xlm/5/6/607/

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580–591. https://doi.org/10.1162/jocn.2008.20054

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience*, *8*, 577. https://doi.org/10.3389/fnhum.2014.00577

van Wijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication*, *35*(1), 103–113. https://doi.org/10.1016/S0167-6393(00)00098-4

Vaughn, C., Baese-Berk, M., & Idemaru, K. (2019). Front & Back Matter. *Phonetica*, *76*. https://www.karger.com/Article/Abstract/502465

Weber, A., & Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *3*(3), 387–401. https://doi.org/10.1002/wcs.1178

Weiss, S., & Mueller, H. M. (2003). The contribution of EEG coherence to the investigation of language. *Brain and Language*, *85*(2), 325–343. https://doi.org/10.1016/s0093-934x(03)00067-1

Wells, J. C. (1982). *Accents of English: Volume 3: Beyond the British Isles*. Cambridge University Press. https://play.google.com/store/books/details?id=484eVQ7t8TMC

Wernicke, C. (1874). *Der aphasische Symptomenkomplex (Breslau, Germany: Cohn and Weigert)*.

Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological Psychology*, *89*(1), 137–147. https://doi.org/10.1016/j.biopsycho.2011.10.002

Wilson, J. P., & Hugenberg, K. (2010). When under threat, we all look the same: Distinctiveness threat induces ingroup homogeneity in face memory. *Journal of Experimental Social Psychology*, *46*(6), 1004–1010. https://doi.org/10.1016/j.jesp.2010.07.005

Winn, M. B., & Moore, A. N. (2018). Pupillometry Reveals That Context Benefit in Speech Perception Can Be Disrupted by Later-Occurring Sounds, Especially in Listeners With Cochlear Implants. *Trends in Hearing*, *22*, 2331216518808962. https://doi.org/10.1177/2331216518808962

Witteman, M. J., Weber, A., & McQueen, J. M. (2014). Tolerance for inconsistency in foreign-accented speech. *Psychonomic Bulletin & Review*, *21*(2), 512–519. https://doi.org/10.3758/s13423-013-0519-8

Wolff, N., Kemter, K., Schweinberger, S. R., & Wiese, H. (2014). What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias. *Social Cognitive and Affective Neuroscience*, *9*(5), 580–590. https://doi.org/10.1093/scan/nst024

Wu, Y.-H., Aksan, N., Rizzo, M., Stangl, E., Zhang, X., & Bentler, R. (2014). Measuring listening effort: driving simulator versus simple dual-task paradigm. *Ear and Hearing*, *35*(6), 623–632. https://doi.org/10.1097/AUD.0000000000000079

Xu, Y., Aubele, J., Vigil, V., Bustamante, A. S., Kim, Y.-S., & Warschauer, M. (2022). Dialogue with a conversational agent promotes children's story comprehension via enhancing engagement. *Child Development*, *93*(2), e149–e167. https://doi.org/10.1111/cdev.13708

Xu, Y., & Warschauer, M. (2020). Wonder with elinor: designing a socially contingent video viewing experience. *Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts*, 251–255. https://doi.org/10.1145/3397617.3398024

Yi, H.-G., Smiljanic, R., & Chandrasekaran, B. (2014). The neural processing of foreign-accented speech and its relationship to listener bias. *Frontiers in Human Neuroscience*, *8*, 768. https://doi.org/10.3389/fnhum.2014.00768

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

9. Appendices

Appendix A: Chapter 3

The list of experimental materials for chapter 3

Afuera

Amor

Azucar

Bajo

Cabeza

Corriendo

Cotilla

Cuento

Deliciosa

Descalzo

Dinero

Manias

Señores

Envidia

Marquesa

Señoritas

Felicidad

Masculinos

Sirvientas

Feliz

Mejor

Sueño

Futbol

Nosotras

Tacita

Gente

Pañuelos

Talones

Inocencia

Pausa

Temores

Jardineros

Princesas

Terrible

Locura

Problemas

Tonterias

Luz

Ropa

Vieja

Appendix B: Chapter 3

Additional Analyses- Time Frequency Analysis

In order to confirm and expand the results of the ERP analysis we also performed a time-frequency power analysis on the EEG data. Time-frequency analysis of EEG single-trial data was done with a Morlet wavelet decomposition using MNE software (Gramfort et al., 2013). This method was used to decompose single-trial time-frequency values between 1 and 40 Hz for the 28 electrodes placed on the scalp (steps = 12). For each accent condition, the resulting power was averaged across trials and electrodes of the cluster. The average total power values were baseline corrected by subtracting the average of the pre-stimulus baseline (− 200 to 0 ms) from each time point separately for every frequency. The resulting change in total power values are displayed in Fig B.1. A one-way repeated measures ANOVA with accent as a three-level factor was run on the average power estimates in the time window of 150 ms to 400 ms (corresponding to the time window where and Accent effect was observed in the ERPs) within the frequency range where the power change was maximum (from frequencies 1 to 5 Hz; see Fig B.2). A main effect of accent was found ($F(2,48) = 3.86$, $p < .05$, $\eta_p^2=0.01$), with a marginally higher power amplitude in the native and dialectal conditions than the foreign condition (Native vs.Foreign:$t(24)=-2.19$,$p=0.06$, Foreign vs.Dialectal: $t(24)=-2.29$, $p=0.06$). There was no significant difference between the native and dialectal power amplitude (Dialectal vs Native: $t(24)=-0.03$, $p=0.97$). After averaging within the above mentioned time window and frequency range, a reduction of power for the foreign condition can be observed. This effect is compatible with what is observed in the ERPs (Fig.B.3.).

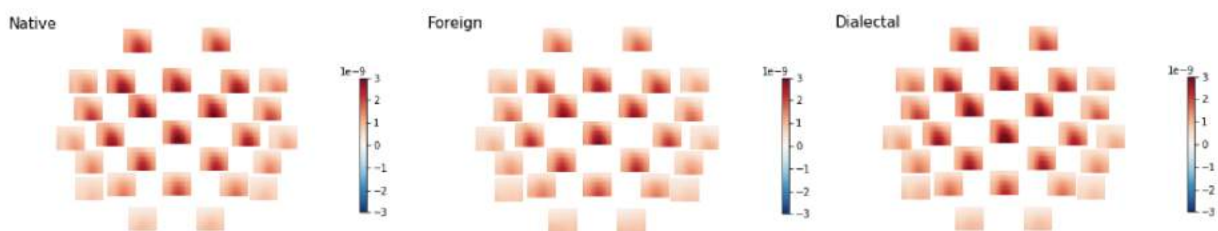


Fig. B.1. Topographic plots of average power amplitudes (1-5 Hz) between 150-400 ms in each accented condition

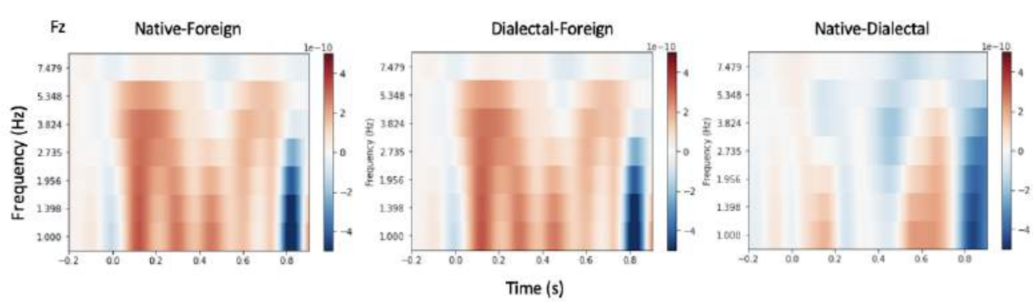


Fig. B.2. Average power amplitude differences between native and foreign, dialectal and foreign and native and dialectal conditions in a representative frontal channel (Fz).

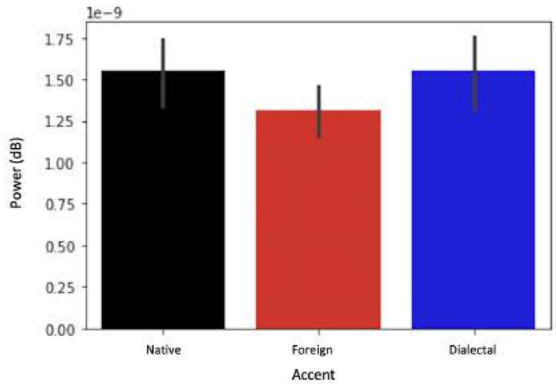


Fig. B.3. Bar plots of the average power amplitudes of each condition (1-5 Hz) in the frontal electrodes between 150-400 ms.

Appendix C: Chapter 4

Overview of selection process for removing incorrectly repeated trials and full dataset averages.

Speech data (sentence repetition) collected using googles automatic speech recognition speech-to-text transcriptions was manually analysed for accuracy. Speech recognition data was used instead of actual recordings due to the change of the experiment from in-lab to online due the pandemic, in order to facilitate data collection before online speech recording ethical protocol was approved at BCBL. We decide to only analyze the data from only correctly repeated trials due to the possible memory effects of not repeating or incorrectly repeating trials and the incompatibility of attempting to link attentional data to memory data if the trials included differ (i.e. if we were to just to remove incorrectly repeated trials from the memory data). If we include all trials in the attentional data as shown below the trends are similar. The speech-to-text transcription included 2 confidence levels in their transcriptions, thus generating to options. In order to decide for accuracy, if the first transcription (always the google model higher confidence version) was correct, we did not look at the second (less confident) transcription. If it was not correct, we first looked at the second transcription and made our judgement considering both. Trials were considered correct if all words or all but one filler word (on, in, a, the) were correct. Incorrect trials we further coded by the reasons they were marked incorrect (e.g. they said I don't know, the didn't speak, they only repeated either the beginning or the end of the sentence, phonological substitution, semantic substitution). Overall, trials were mostly correctly repeated (average percentage correctly repeated: Dialectal: 87% Native: 88% Foreign: 80%).

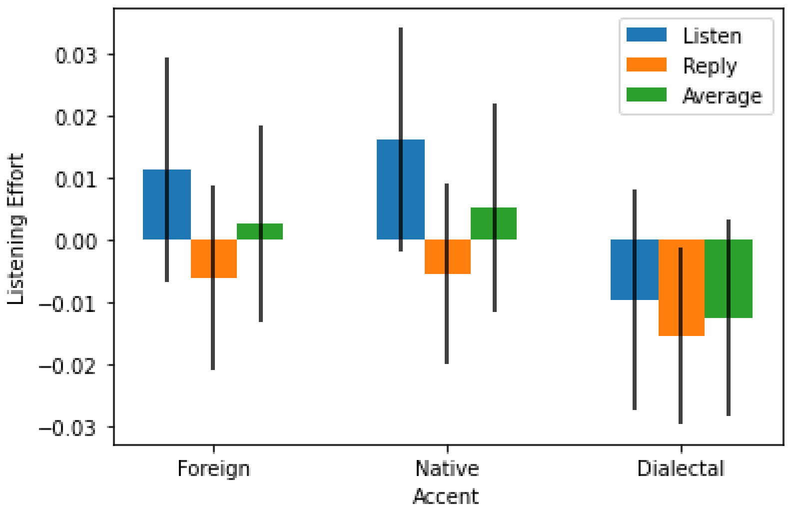
Attentional data with all trials (correct and incorrectly repeated) included.

Raw Time-on-Target Data

		Foreign			**Native**			**Dialectal**		
	RPM	**LISTEN**	**REPLY**	**AVG**	**LISTEN**	**REPLY**	**AVG**	**LISTEN**	**REPLY**	**AVG**
average	3.3	77.7	79.2	78.4	77.5	79.2	78.4	79.5	79.9	79.7
sd	1.4	6.7	7.6	6.8	7.0	6.6	6.5	6.5	6.4	6.2
sterr	0.2	0.9	1.1	1.0	1.0	0.9	0.9	0.9	0.9	0.9

Objective Listening Effort Scores of all trials, corrected by the End Baseline

	Foreign			**Native**			**Dialectal**		
	LISTEN	**REPLY**	**AVG**	**LISTEN**	**REPLY**	**AVG**	**LISTEN**	**REPLY**	**AVG**
average	0.01135305	-0.006121	0.00261601	0.0159841	-0.0054873	0.00524842	-0.0097627	-0.0154976	-0.0126301
sd	0.12145221	0.12150998	0.1185855	0.09976389	0.09810018	0.09535024	0.1062134	0.11269656	0.10690986
sterr	0.01810503	0.01811364	0.01767768	0.01487192	0.01462391	0.01421397	0.01583336	0.01679981	0.01593718



Appendix D: Chapter 4

List of old and new sentences used for the memory experiment.

"Old" Sentences
(Spanish)

Para entrar en la élite hace falta un permiso.
Hacía sol pero una nube empezó a acercarse.
Necesitaba una inyección de insulina.
Solo hizo una corrección en el texto del alumno.
Sin darme cuenta, la miel se acabó.
El pasado semestre, el examen que hice fue complicado.
Me di un golpe en la rodilla y todavía me duele.
En la generación de mis abuelos se tenían más hijos.
Afortunadamente, la producción de sidra está aumentando.
Prefiero el cojín muy blando.
Se puso el delantal antes de empezar a cocinar.
Ayer el cable de la radio se estropeó.
Creo que el maquillaje de la señora era exagerado.
Estoy hirviendo una coliflor del Eroski.
Para cuidar la mente es importante hacer ejercicio.
Intentó adivinar la clave pero no pudo.
Llegué hasta la fuente y me bañé.
Si tuviera que elegir, el deporte vasco que más me gusta es la pelota.
En verano el jardín se seca.
Alguien miró el bosque desde la ventana.
En Pasaia, la construcción del barco está tardando un poco.
Como dice el refrán, la parte más dura de la cuesta tiene que llegar.
No puedes imaginar el hambre que tengo.
Vi que el barril de sidra se estaba acabando.
Me contaron un chiste que no sabía.
Entre la gente había absoluto silencio.
Supe que la calle era conocida de varias formas.
Al final la gestión de la empresa llegó a ser problemática.
Vi que la rebelión del hijo no tardó en llegar.
Parece que el crimen del ladrón fue rápido.
Dijo que la celebración de San Fermín es en Julio.
Saqué el mantel de la cesta.
No encontró la dirección del restaurante.

Nunca puedo saltarme el postre de queso.
Rellenó el informe con todos los detalles.
Compré un rape para poner al horno.
Tengo un lunar en la pierna.
Sé que el brócoli de esta tienda es muy sabroso.

(English)
To enter the elite you need a permit.
It was sunny but a cloud began to approach.
I needed an insulin injection.
She only made a correction to the student's text.
Without realizing it, the honey ran out.
Last semester, the exam I took was complicated.
I hit my knee and it still hurts.
In my grandparents' generation there were more children.
Fortunately, cider production is increasing.
I prefer the very soft cushion.
She put on her apron before starting to cook.
Yesterday the radio cable broke.
I think the lady's makeup was overdone.
I am boiling a cauliflower from the Eroski.
To take care of the mind it is important to exercise.
She tried to guess the password but she couldn't.
I got to the fountain and bathed.
If I had to choose, the Basque sport that I like the most is pelota.
In summer the garden dries up.
Someone looked at the forest from the window.
In Pasaia, the construction of the ship is taking a while.
As the saying goes, the hardest part of the climb has to come.
You can't imagine how hungry I am.
I saw that the barrel of cider was running out.
They told me a joke I didn't know.
Among the people there was absolute silence.
I learned that the street was known in various ways.
In the end the management of the company became problematic.
I saw that the son's rebellion was not long in coming.
It seems that the thief's crime was quick.
She said that the celebration of San Fermín is in July.
I took the tablecloth out of the basket.
He couldn't find the address of the restaurant.

I can never skip the cheese dessert.
She filled out the report with all the details.
I bought a monkfish to put in the oven.
I have a mole on my leg.
I know that the broccoli in this store is very tasty.

"New" Sentences
(Spanish)

Para entrar en la cueva hace falta un abrigo.
Hacía sol cuando una nube empezó a alejarse.
Necesitaba una estación de repostaje.
Solo hizo una anotación en el libro del alumno.
Sin darme cuenta, la nuez se pudrió.
El pasado octubre, el proyecto que hice fue complicado.
Me dió un golpe en la costilla y todavía me duele.
En la generación de mis alumnos se tenían más libros.
Afortunadamente, la empresa de sidra está triunfando.
Prefiero el cerdo muy gordo.
Se puso el vestido antes de empezar a trabajar.
Ayer el botón de la radio se atascó.
Creo que el mayordomo de la señora era americano.
Estoy cocinando una borraja del Eroski.
Para cuidar la línea es importante hacer régimen.
Intentó adivinar la letra pero no veía.
Llegué hasta la calle y me paré.
Si tuviera que elegir, el pintxito vasco que más me gusta es la kokotxa.
En Granada el pulpo se seca.
Alguien tiró el gatito desde la ventana.
En Pasaia, la construcción del piso está tardando un montón.
Como dice el refrán, la parte más larga de la cuesta tiene que costar.
No puedes entender el problema que tengo.
Vi que el vaso de sidra se estaba rompiendo.
Me contaron un rumor que no creía.
Contra la norma había absoluto silencio.
Dije que la chica era conocida de varias formas.
Al final la reunión de la empresa llegó a ser irrelevante.
Vi que la herencia del padre no tardó en llegar.
Parece que el viaje del cura fue rápido.
Dijo que la celebración de San Marcial es en junio.

Saqué el botón de la blusa.
No encontró la lámpara del vestíbulo.
Nunca puedo dibujar el trozo de queso.
Rellenó el examen con todos las respuestas.
Compré un pollo para hacer al ajillo.
Tengo un primo en la fiesta.
Sé que el interior de esta tienda es muy lujoso.

(English)

To enter the cave you need a coat.
It was sunny when a cloud began to move away.
I needed a refueling station.
He only made an entry in the student's book.
Without realizing it, the nut rotted.
Last October, the project I did was complicated.
He hit me in the rib and it still hurts.
In the generation of my students there were more books.
Fortunately, the cider company is succeeding.
I prefer very fat pig.
He put on the dress before starting work.
Yesterday the radio button got stuck.
I think the lady's butler was American.
I'm cooking a borage from Eroski.
To take care of the line it is important to diet.
She tried to guess the letter but she didn't see.
I got to the street and stopped.
If I had to choose, the Basque pintxito that I like the most is the kokotxa.
In Granada the octopus dries up.
Someone threw the kitten from the window.
In Pasaia, the construction of the apartment is taking a long time.
As the saying goes, the longest part of the slope has to cost.
You can't understand the problem I have.
I saw that the glass of cider was breaking.
They told me a rumor that I didn't believe.
Against the norm there was absolute silence.
I said that the girl was known in various ways.
In the end the company meeting became irrelevant.
I saw that the father's inheritance was not long in coming.
It seems that the journey of the priest was fast.

He said that the celebration of San Marcial is in June.
I took the button off the blouse.
He didn't find the hall lamp.
I can never draw the piece of cheese.
She filled out the test with all the answers.
I bought a chicken to make garlic.
I have a cousin at the party.
I know that the interior of this store is very luxurious.

Appendix E: Chapter 4

Additional analyses using the beginning and averaged beginning and ending baseline measures.

1. Listening Effort Analyses with Additional Baseline Measures

Baseline Start

A one way ANOVA with accent as a three level factor (native, dialectal, foreign) was run on the averaged listening effort scores that were corrected by the starting baseline measure. Similarly to the end baseline used in the chapter, no significant effect of accent was found (($F(2,90)=1.8$, $p = 0.17$).

Averaged Baseline

A one way ANOVA with accent as a three level factor (native, dialectal, foreign) was run on the averaged listening effort scores that were corrected by averaging the beginning and ending baseline measures. Similarly to the end baseline used in the chapter, no significant effect of accent was found (($F(2,90)=2.2$, $p = 0.11$).

2. Listening Effort Analysis with only 31 participants who completed subjective listening effort measures included

A one way ANOVA with accent as a three level factor (native, dialectal, foreign) was run on the averaged listening effort scores that were corrected by the baseline. Similarly to the whole data sample, no significant effect of accent was found (($F(2,90)=1.05$, $p = 0.35$). Exploratory follow-up T-tests revealed a marginally significant difference between dialectal and foreign accent, where foreign accent was associated with greater listening effort than dialectal accent ($t(30)=1.96$, $p=0.06$). Listening effort of foreign accent was also marginally higher than native accent ($t(30)=1.80$, $p=0.08$). No significant difference was found between dialectal and native listening effort ($t(30)=0.99$, $p=0.33$). As these results are very similar to the ones of the full sample we believe that correlations between listening effort and memory can be similarly made for both the entire sample and for the sample that we have the subjective listening effort measure of.

Appendix F: Chapter 4

1. Details on 'remember' 'know' classification of the memory trials

After responding whether the sentence that appeared on the screen was old (they heard it in the attentional part of the experiment) or new (they did not listen to it before), participants were asked to respond with key presses categoring their memory into 'know' (it seems familiar to me but I cannot recall explicitly when I heard it) and 'remember' (I explicitly remember hearing this sentence). After that they were also asked what accent they thought they heard the trial in.

Average percentage of 'hit' trials (correctly identified as old/new) that were know/remember.
0.244895293 (sd: 0.18) Know
0.755104707 (sd: 0.18) Remember

Most hits were trials that the participants marked as 'remember' trials, supporting the recognition memory literature that recollection memory trials seem to be more deeply encoded than familiarity trials.

Average percentage of correct identification of the accent of each 'hit' trial
0.455023449 (sd: 0.22) foreign accent
0.660887646 (sd: 0.21) native accent
0.303092432 (sd:0.21) dialectal accent

Participants still struggled even in 'remember' trials that they got correct to identify the speaker accent that they heard the trial in. However, they were best at correctly remembering when it was heard in their native accent.

2. Survey Questions

(Spanish)

¿Consideras el acento guipuzcoano como tu acento nativo?
¿Te consideras bilingüe? (¿hablas más de un idioma con fluidez?)
¿En tu opinión, qué acento fue más atractivo o interesante de escuchar?
¿Crees que repetir las frases hizo más fácil o más difícil recordar la frase más adelante en el experimento?
¿Algún acento fue más difícil de repetir que los demás?
¿En una escala del 1 a 10 (1-muy fácil, 10-muy difícil) como de difícil era prestar atención el acento inglés?
¿En una escala del 1 a 10 (1-muy fácil, 10-muy difícil) como de difícil era prestar atención el

acento argentino?
¿En una escala del 1 a 10 (1-muy fácil, 10-muy difícil) como de difícil era prestar atención el acento vasco?
¿En una escala del 1 a 10 (1-muy fácil, 10-muy difícil) como de difícil era memorizar las frases que escuchaste en el acento inglés?
¿En una escala del 1 a 10 (1-muy fácil, 10-muy difícil) como de difícil era memorizar las frases que escuchaste en el acento argentino?
¿En una escala del 1 a 10 (1-muy fácil, 10-muy difícil) como de difícil era memorizar las frases que escuchaste en el acento vasco?
¿Crees que repetir las frases hizo más fácil o más difícil recordar la frase más adelante en el experimento?
¿Algún acento fue más difícil de repetir que los demás?
¿Sabes inglés? ¿Qué nivel tienes?
¿Has estado en Inglaterra, Reino Unido, EE.UU. O Canadá durante más de dos semanas?
¿Eres originario de la Comunidad Autónoma Vasca? ¿Y tu familia?
¿Tienes amigos, compañeros de trabajo, compañeros de clase ingleses, de EE.UU. o Canadá con los que interactúes de forma regular?
¿Tienes amigos, compañeros de trabajo o compañeros de clase argentinos con los que interactúes de forma regular?
¿Cuántas horas a la semana estás expuesto a un acento extranjero (no nativo en español)?
¿Cuántas horas a la semana estás expuesto al acento de un hablante nativo inglés?
¿Cuántas horas a la semana estás expuesto al acento argentino o de Uruguay?
¿Era más difícil seguir el círculo mientras haciendo la tarea de escuchar y repetir frases?
¿Qué dispositivos has utilizado para realizar el experimento?

(English)
Do you consider the Gipuzkoan accent as your native accent?
Do you consider yourself bilingual? (do you speak more than one language fluently?)
In your opinion, which accent was more attractive or interesting to listen to?
Do you think repeating the phrases made it easier or harder to remember the phrase later in the experiment?
Was any accent more difficult to repeat than the others?
On a scale of 1 to 10 (1-very easy, 10-very difficult) how difficult was it to pay attention to the English accent?
On a scale from 1 to 10 (1-very easy, 10-very difficult) how difficult was it to pay attention to the Argentine accent?
On a scale from 1 to 10 (1-very easy, 10-very difficult) how difficult was it to pay attention to the Basque accent?
On a scale of 1 to 10 (1-very easy, 10-very difficult) how difficult was it to memorize the phrases you heard in the English accent?

On a scale from 1 to 10 (1-very easy, 10-very difficult) how difficult was it to memorize the phrases you heard in the Argentine accent?
On a scale from 1 to 10 (1-very easy, 10-very difficult) how difficult was it to memorize the phrases you heard in the Basque accent?
Do you think repeating the phrases made it easier or harder to remember the phrase later in the experiment?
Was any accent more difficult to repeat than the others?
Do you know English? What level do you have?
Have you been in England, UK, USA or Canada for more than two weeks?
Are you from the Basque Autonomous Community? And your family?
Do you have English, US or Canadian friends, co-workers, classmates that you interact with on a regular basis?
Do you have Argentinian friends, co-workers or classmates that you interact with on a regular basis?
How many hours a week are you exposed to a foreign accent (not native in Spanish)?
How many hours a week are you exposed to the accent of a native English speaker?
How many hours a week are you exposed to an Argentine or Uruguayan accent?
Was it more difficult to follow the circle while doing the task of listening and repeating sentences?
What devices did you use to carry out the experiment?

Appendix G: Chapter 5

List of Comprehension Questions

(Spanish)
1.¿Durante qué momento del día transcurre el diálogo?
2. ¿Qué está intentando abrir Emma?
3. ¿Qué oficios tienen los personajes?
4. ¿A quién quiere parecerse Emma?
5. ¿Qué tipo de temperatura hace en el exterior?
6. ¿Quién había avisado a Sara de que desconfiase de cierta persona?
7. ¿Cuál es la palabra que empieza por R y que le da miedo a Sara?
8. Mientras hablan ¿está la lámpara encendida?
9. ¿Qué bordan las señoritas de la casa?
10. ¿Qué desayunan los personajes?

(English)
1. During what time of day does the dialogue take place?
2. What is Emma trying to open?
3. What jobs do the characters have?
4. Who does Emma want to seem like?
5. What temperature is it outside?
6. Who had warned Sara that she should distrust a certain person?
7. What is the word that begins with R and that scares Sara?
8. While they are talking, is the lamp on?
9. What do the ladies of the house embroider?
10. What do the characters eat for breakfast?